The Imitation Game: Polanyi vs. Turing and Why it Matters to Human Dignity

Andy Steiger

Introduction: Technological Advancement and Human Dignity

Given the rapid technological growth of Artificial Intelligence (AI) through advancements in machine learning, what was once thought impossible is quickly becoming a reality. It is no longer so far-fetched that human-like machines will soon be a part of everyday life. People today are divided whether these continued advancements in AI technologies will lead to the best of times or the worse of times for mankind. Historical studies and physiological research continue to demonstrate the dangers of dehumanization. Given AI's ability to dehumanize on multiple levels, if we are to avoid the mistakes of our past, our future will depend on people's ability to correctly see their humanity and the humanity of others.

By developing the work of Michael Polanyi and Alan Turing, the following essay juxtaposes reductionist perspectives in AI studies that are dehumanizing and explores an alternative foundation that can help navigate the technological future, while upholding inherent dignity. To accomplish this, the following article is broken into four parts: first, a brief history of AI's ontological development is developed within Polanyi and Turing's interactions through tacit knowledge and the imitation game; second, it is demonstrated that Turing's imitation game, as expressed in *strong AI*, undermines the inherent nature of dignity in that it is intrinsically dehumanizing; third, Polanyi's machine ontology is developed as an alternative to the imitation game; fourth, Polanyi's, non-reductive approach is contrasted with Turing's reductive approach, to explore which best provides a foundation for the nature of human dignity.

1. Artificial Intelligence – Polanyi vs. Turing Machine Ontology

For a time Michael Polanyi and Alan Turing both worked at Manchester University in the UK. Polanyi was a scientist-turned-philosopher and Turing was a mathematician-turned-computer scientist. They both had a keen interest in AI studies and regularly discussed the philosophy of "Intelligent Machines" (Andrew Hodges, 2009, p. 13). Turing was focused on technological advancement and Polanyi was concerned with the philosophies behind these advancements. Through their conversations Polanyi encouraged Turing to publish his views and organized formal discussions on the subject. On October 27, 1949, a significant interdisciplinary panel discussion was held on "The Mind and the Computing Machine" consisting of Michael Polanyi, Alan Turing, M. Newman, M. B. Bartlett, J. Z. Young, D.N. Emmett and others to discuss concerns with the philosophy behind AI (William Taussig Scott & Martin X. Moleski, 2005, p. 215).

The key question driving that dialogue was: *Can thinking be mechanical*? Although there is no complete transcript from that dialogue, a rough outline of the proceedings has survived, in which Polanyi is recorded as having argued against mechanical thinking given that the 'semantic function, [is] outside the formalizable system' (*Discussion on The Mind and the Computing Machine*, 1949). In reference to cybernetics, Polanyi explains what he means by the semantic function:

> I maintain that a formal system of symbols and operations functions as a deductive system only by virtue of unformalised supplements. We must know the meaning of undefined terms, understand what is stated in our axioms and believe it to be true, and acknowledge an implication in the handling of symbols by formal proof. These acts of

> knowing, understanding and acknowledging are not formalised: they may be jointly designated as the 'semantic operations' of the formalised system.

(Michael Polanyi, 1952, p. 313)

This, Polanyi argued, was definitively demonstrated by Gödel through his *incompleteness theorem* in that even mathematics cannot be verified within its own system but requires persons capable of working outside the system (Michael Polanyi, 1962, pp. 260–261). Polanyi's point was to emphasize the importance of personal knowledge, referring to these 'unformalized operations' as tacit knowledge (commonly known today as *Polanyi's Paradox*), which is encapsulated in his famous axiom: "We can know more than we can tell" (Michael Polanyi, 1966, p. 4). Specifically, Polanyi argued that the mind operated according to tacit knowledge or unspecified knowledge, which made things like playing chess or diagnosing a disease impossible to code into a machine, given these cannot be completely specified in an explicitly step-by-step fashion to be coded for.

Realizing the difficulties raised by Polanyi and others, Turing proposed a creative alternative. The following year, in 1950, Turing published his famous paper *Computing Machinery and Intelligence,* in which he writes,

> Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words. The new form of the problem can be described in terms of a game which we call the 'imitation game.'

(A. M. Turing, 1950, p. 433)

The imitation game is based on deception in which an interrogator, person A (either male or female) is in an isolated room and holds a text-only conversation with person B (male) and person C (female). The goal of the game is for person B to deceive the interrogator into thinking that he is a girl and for person C to convince the interrogator that she is a girl. At the end of the game the interrogator (person A) must correctly identify the gender of both B and C. If B or C are successful in either deceiving or convincing the interrogator—they win.

Turing proposed a version of this game in which either person B or C is swapped with a machine. The question Turing asked was, "Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman?" (A. M. Turing, 1950, p. 433). There are many variations of this game that are commonly known today as the Turing Test. The premise of the Turing Test has remained the same: instead of defining thinking, a machine is programed to imitate the behaviour of thinking, which Turing believed qualified as thinking and avoided the paradox of an explicit definition.

It is helpful to note that the imitation game was not a new concept; the desire to create an android, (Greek for "manlike") was coined by Gabriel Naudé in the 17th century. By the 18th century, people already began to create manlike machines that mimicked their human creators, from Vaucanson's flute playing android to Kempelen's chess playing machine called The Turk. In fact, these machines led to new discoveries. For example, Vaucanson's machine shed new insights into music theory and The Turk, which was exhibited across Europe and America competing against and beating both Napoleon and Benjamin Franklin, demonstrated people's interest in these machines and the complexities of creating a thinking machine (Jessica Riskin, 2016, pp. 115–125). The Turk was exposed as a fake, just a small man cleverly hidden in a box. Over the centuries, the continued interest in these manlike machines has led to technological advancements, such as the universal Turing Machine, in which it is no longer necessary to hide a man in the box, but uses cleverly devised algorithms, which has raised the philosophical discussion concerning the nature of these machines.

Challenged by Turing and the advancement of computers, Polanyi still affirmed his argument, even given the potential of new technological advancements he called 'automatic operations' explaining,

> The proliferation of axioms discovered by Gödel offers manifest proof that a person operating a logical inference machine can achieve informally a range of knowledge which no operations of such a machine can demonstrate, even though its operations suggest an easy access to it. It proves that the power of the mind exceed those of a logical inference machine. But we have yet to face the wider problem raised by gunsight predictors, automatic pilots, etc., that is, by machines whose performances range far beyond logical inferences. A. M. Turing has shown that it is possible to devise a machine which will both construct and assert as new axioms an indefinite sequence of Gödelian sentences. Any heuristic process of a routine character—for which in the deductive sciences the Gödelian process is an example—could likewise be carried out automatically. A routine game of chess can be played automatically by a machine, and indeed, all arts can be performed automatically to the extent to which the rules of the art can be specified. While such a specification may include random elements, like choices made by spinning a coin, no unspecifiable skill or connoisseurship can be fed into a machine'.
>
> (Michael Polanyi, 1962, p. 261, nn 2)

Polanyi was correct, computer science has shown that it is impossible to code a machine to play chess at a high level through explicit step-by-step instructions but he failed to consider that a machine could be coded with algorithms that learn to play chess at a high level. Technologies, such as machine-learning algorithms, have now advanced computers beyond their human creators. For example, in 1996, IBM's computer *Deep Blue* defeated the world champion chess player Gary Kasparov. In 2011, IBM's machine *Watson* defeated two of Jeopardy's greatest players. More recently, in 2016, Googles *AlphaGo* algorithm beat the world *Go* champion Sedol Lee 4-1, a feat that many thought either impossible or decades from being a reality. The early win was made possible through advances in deep learning and reinforcement learning that allowed the algorithm to compete against itself, which has now advanced to *AlphaGo Zero,* which is capable of learning the game through playing alone and has gone on to defeat the original *AlphaGo* program 100-0 (David Silver et al., 2017, p. 354).

Machine learning has overcome aspects of Polanyi's Paradox through the use of what can be called tacit algorithms such as a confusion matrix, Monte Carlo algorithms, and neural networks. Inspired by brain neurons, neural networks process large quantities of data through multiple layers called deep learning so as to fine tune hidden variables that act as a type of weak tacit knowledge. These tacit algorithms have achieved a new level of statistical precision that allow machines to learn more than just games. For example, Polanyi was a physician before going into physical chemistry and from his experience believed it impossible for a machine to accurately diagnose a patient's disease given the high level of skill and tacit variables involved (Michael Polanyi, 1969a, p. 132). However, machine learning algorithms have even demonstrated this level of tacit ability, having an accuracy of determining skin cancer on par with that of a dermatologist (Andre Esteva et al., 2017, pp. 115–118). Polanyi believed that a machine could not learn these skills because they cannot be communicated in a step-by-step process but require tacit learning through apprenticeship. However, recent advancements in 'deep feature learning' have shown even this is possible; machines can imitate 'component[s] of human expertise learning' (Nan Li, Noboru Matsuda, William W. Cohen, & Kenneth R. Koedinger, 2015).

Turing envisioned the day when computers would learn to play Chess or Go in a human vs. machine match, but the ultimate challenge would be the imitation game. (A. M. Turing, 1996, p. 257). Given the advancements in machine learning, there is little doubt that one day a machine will pass at least a simple form of the Turing Test (such as a text only conversation) and potentially one day pass what is called the total Turing Test (speech and a face-to-face conversation). Given continued technological advancements, the important question that must be considered is not *if* a machine will win the imitation game but *what* the prize will be.

Within Computer Science two schools of AI have emerged, a weak (cautious or narrow) version and a strong version. Searle coined this distinction, defining weak AI as "the principal value of the computer in the study of the mind"; whereas, according to strong AI,

> the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really *is* a mind in the sense that computers given the right programs can be literally said to *understand* and have other cognitive states. In strong AI, because the programmed computer has cognitive states, the programs are not mere tools that enable us to test psychological explanations; rather, the programs are themselves the explanations.
>
> (John R. Searle, 1980, p. 417.)

Following these distinctions, passing the imitation game, according to Searle can be interpreted as a weak Turing Test, replicating human behaviour, and a strong Turing Test, the human mind is behaviour, ((John R. Searle, 2009, p. 141). Thus, strong AI is commonly associated with the strong Turing Test, becoming more than a game but an ontological test of humanity that Polanyi acknowledged and challenged (Michael Polanyi, 1964, 84-85.) and Turing embraced and believed would one day go beyond humanity (A. M. Turing, 1996, pp. 259–260).

On July 12, 2017, an indication of how culture views AI and what the prize for winning the imitation game might be was demonstrated when a robot named Sophia was granted citizenship to Saudi Arabia during the tech summit at the Future Investment

Initiative. This lifelike machine created by Hanson Robotics is designed to look and act human, everything from skin tone, facial expressions, and language processing capabilities. Sophia represents more than another step forward in humanity's drive toward creating innovative technologies but also represents people's growing propensity to anthropomorphize machines. Consider that although Sophia could not pass either a weak or strong version of the Turing Test, it was still granted citizenship. This not only demonstrates people's ability but also their desire to humanize machines and reveals a human bent toward embracing strong AI. If caution is not taken this could include more than citizenship to a country but to the human race with all the benefits that come with being a person, as indicated in the Universal Declaration of Human Rights (UDHR) such as inherent dignity. However, I will argue that following the ethic of the UDHR, granting machines human status does not bring machines to the level of humanity, it does the exact opposite—it brings humans down to the level of a machine.

2. Dehumanization – The Danger of Denying and Imitating Dignity

Dehumanization is the very reason that the UDHR was created after WWII. It was a response to such things as eugenics and the atrocities that followed, such as the genocide of the Jews, when persons no longer saw each other as fully human. Historically, people have tended not to question if humans have dignity, rather they have questioned who qualifies as human, which is at the heart of the dehumanizing nature of strong AI. When people attempt to humanize a robot, by placing it on par with humanity, it has the exact opposite effect; it dehumanizes people by lowering them to that of a machine. The reason for this is implicit in the UDHR.

The UDHR clearly begins by stating that humans have *"inherent* dignity," meaning that dignity is not bestowed by humans—it is recognized; that is, humans are born with dignity (*A Universal Declaration of Human Rights*, n.d.). However, when people attempt to bring a machine to the level of a human, it undermines humanity's dignity because it is no longer inherent. The machine earns its humanity, for example, by passing the imitation game. Granting a machine humanity through something like the Turing Test makes dignity no longer inherent but conditional and herein lies the problem. Depending on how narrow or loose one draws humanity's lines of demarcation, some people will naturally be included and others excluded. For example, if humanity is conferred by convincingly imitating a person, what of those humans, such as children or the disabled, that cannot pass the test? There are many humans unable to communicate sufficiently to pass the Turing Test, in which case this test would grant humanity to some machines and remove it from some humans. Peter Singer uses a similar line of reasoning when he argues against speciesism among animals. He writes:

> [T]here will surely be some nonhuman animals, whose lives, by any standards, are more valuable than the lives of some humans. A chimpanzee, dog, or pig, for instance, will have a higher degree of self-awareness and a greater capacity for meaningful relations with others than a severely retarded infant or someone in a state of senility. So, if we base the right to life on these characteristics we must grant these animals a right to life as good as, or better than, such retarded or senile human beings.
>
> (Peter Singer, 1975, p. 19)

What Singer is advocating for here is an imitation game played with animals. Consider that some humans, such as a child or the disabled are considerably easier for a machine or animal to imitate. But this raises the question: What is the standard of human behaviour or, in this case, self-awareness that must be mimicked? As well, there are those that will be more demanding interrogators or judges of the Turing Test than others, as demonstrated with Sophia. Who decides what passes as human behaviour? This can be challenging given a simple flaw within the Turing Test, such as that demonstrated by Warwick and Shah in asking,

> [I]f the most appropriate response, (as deemed by an entity) to a particular question, is silence? What is the thinking nature of a machine when, rather than responding to an inappropriate or inane question, it does not answer it with an utterance? Why should a truly intelligent machine ingratiate itself with humanlike responses just to be considered human? Is not the truly Turing-intelligent machine the one that knows when and why to be silent?
>
> (Kevin Warwick & Huma Shah, 2017, p. 287)?

It is conceivable that a person or machine could employee the simple strategy of pleading the 5th, such as in a court of law, and thereby potentially pass the Turing Test.

As technology advances and increasingly becomes more proficient at mimicking its human creators, caution needs to be taken to avoid the dehumanizing mistakes of the past. It has been recognized that dehumanization, which robs a person of their dignity, takes place in one of two ways: by denying a person their human nature*,* equating them as an object, or denying their *uniqueness* equating them as an animal (Nick Haslam, 2006, p. 256). As well, there is another form of dehumanization that can be understood as dementalization, being the denial of mind, which does not necessarily constitute a third category but as Haslam explains, "denial of mind implies a denial of humanness on both of the dimensions." (Nick Haslam, 2014, p. 37).

All three forms of dehumanization are present in strong AI. First, as strong AI attempts to humanize machines, it simultaneously dehumanizes people because it is predicated on the assumption of what Polanyi called the "modern scientific outlook" of reductive physicalism. Polanyi, borrowing from Buber, explains that reductive physicalism "tends towards replacing everywhere the personal I-Thou by an impersonal I-It"(Michael Polanyi, 1957, p. 331). Accordingly, it robs humanity of their *nature* by reducing persons to their biological and physical parts operating according to DNA's code and nature's laws. Second, this also leads to the denial of human *uniqueness* as people are equated to biological machines, an animal made of similar parts but following a different biological code. Thus, the only difference between humans, animals, and machines on reductionism is their behaviour. Third, human nature and uniqueness are both denied by reducing the mind or consciousness to biological brain states, and further by equating brain states as either random events of quantum indeterminacy or determined actions of particulars following the laws of Newtonian physics.

Within this three-part framework, the imitation game has introduced a developing implication of dehumanization that is beginning to be studied and can be understood as artificial dignity (AD). For example, it can be argued that a machine that passes a weak Turing Test has a weak level of AD and passing a strong Turing Test has a high level of AD. This is *not* to say that a machine has real dignity, but the image of humanity that it is imitating has dignity and as such, it has a dehumanizing potential for one's self and society. In robot ethics, this is similar to what Danaher refers to as the "symbolic-consequences argument" (John Danaher, 2017, p. 107). Danaher demonstrates that symbolic-consequences can be argued at a very basic level with words that reflect an offensive meaning or at a more complex level with machines, such as when an android is used to perform objectionable acts (John Danaher, 2017, pp. 113–114). Robot sex is a strong example of the consequences of AD and thus will be used to make this point.

Given people's ability to now create humanlike machines for sexual purposes, people can abuse those anthropomorphized relationships. For example, Gutiu writes:

> To the user, the sex robot looks and feels like a real woman who is programmed into submission and which functions as a tool for sexual purposes. The sex robot is an ever-consenting sexual partner and the user has full control of the robot and the sexual interaction. By circumventing any need for consent, sex robots eliminate the need for communication, mutual respect, and compromise in the sexual relationship. The use of sex robots results in the dehumanization of sex and intimacy by allowing users to physically act out rape fantasies and confirm rape myths.
>
> (Sinziana Gutiu, n.d., p. 2)

A gynoid designed to imitate rape fantasies is dehumanizing but it is not the machine that is being denied its humanity, as it has no humanity to deny, rather the human user is dehumanized by conditioning their view of themselves and other human beings. Gutiu makes a similar point explaining, "Sex robots cause harm because they provide the user with an illusion of a mutual sexual experience, while also further alienating them from society and normalizing dehumanization of women" (Sinziana Gutiu, n.d., p. 15).

This is the dark side of the imitation game; a sex robot behaves according to the desire of its creator, which raises the concern that this will influence the way that users see and treat the dignity of real people. For example, it has been shown that, "child pornography offending is a valid diagnostic indicator of pedophilia" (Michael C. Seto & James M. Cantor, 2006, p. 613). That being the case, upholding human dignity requires society to question if people should be allowed to make virtual child pornography? In April 2002, the United States Supreme Court found the Child Pornography Prevention Act unconstitutional. Although making and showing sexually explicit pictures of real children is illegal, the manufacture of virtual images is permitted within two categories:

(a) Sexually explicit pictures of actual models who appear to be younger than they are, and
(b) Computer-generated sexually explicit pictures of children.

Virtual pornography was permitted by the courts because it was determined not to harm a specific child. However, the courts neglected to see the harm caused to the user by their actions and the potential harm to those in society that are being dehumanized within the user's worldview.

As technology has moved from virtual pornography to physical pornography, it now must be again asked if companies should be allowed to make machines in the image and likeness of a child for sexual purposes? The U.S. Congress passed the CREEPER bill in 2017 (Curbing Realistic Exploitative Electronic Pedophilic Robots). "Specifically, the bill makes it a crime to import, or knowingly use a common carrier or interactive computer service to transport in interstate or foreign commerce, a child sex doll" (CREEPER Act of 2017, n.d.). Whether or not this bill will be passed into law has yet to be seen but in the UK, child sex dolls have led to arrests and imprisonment for, "importing an obscene article ("Andrew Dobson jailed for 'child-like' sex doll import bid," 2017).

The dangers of dehumanization are clearly documented historically and research continues to show that, "Moral action and moral judgment seem to depend on a appreciation of the humanness of others." (Nick Haslam, Brock Bastian, Simon Laham, & Stephen Loughnan, 2012, 203). Given the relative ease with which people can rob another of their humanity virtually and physically and the terrible results that follow in the real world, it is of the utmost importance that dehumanization continue to be addressed. This is especially concerning given recent technological advancements in AI and robotics, which makes it more important than ever to understood why a machine is not human and what a machine is.

3. Ontology – What is a machine?

Polanyi was adamant that a machine was not human. He distinguishes between machines and humans through a commitment to a non-reductive understanding of persons that he called personal knowledge. Highlighting this personal nature of humanity *not* explained via a reductive ontology, Polanyi writes:

> Our existing knowledge of physics and chemistry can certainly not suffice to account for our experience of active, resourceful living beings, for their activities are often accompanied by conscious efforts and feelings of which our physics and chemistry know nothing. But let us assume for the sake of argument that physics and chemistry could be expanded to account for the sentience of certain physico-chemical systems. It might not be inconceivable that a machine of sufficient complexity would develop conscious thinking, without losing its machine-like character. However, conceived in this sense, conscious thoughts would be the mere accompaniment of automatic operations, on the outcome of which they could exercise no influence.
>
> (Michael Polanyi, 1962, p. 336)

Here Polanyi begins to highlight the mind-body problem by differentiating between what can be understood as weak and strong forms of tacit knowledge. A weak form of tacit knowledge is demonstrated in detached machine learning that imitates the operations of the brain, whereas strong tacit knowledge is personal in nature, in that the mind has intentionality. Thus, although machine learning is inspired by humans, it is ontologically distinct and to treat it as the same would be mistaken. Alpaydin notes that an ontology of imitation is not used outside of AI studies; referring to machine learning he writes, "Our immediate source of inspiration is the human brain, just as birds were the source of inspiration in our early attempts to fly…nowadays, we see birds and airplanes as two different ways of flying—we call them airplanes now, not artificial birds" (Ethem Alpaydin, 2016, pp. 85–86). No one seriously suggests that an airplane is a bird because it employs aeronautics. In the same way, a machine should not be considered human because it employs neural networks to accomplish tacit feats of intelligence.

Going even further, Wittgenstein articulated the impossibility of attaining this strong form of tacit knowing by explaining, "If a lion could talk, we could not understand him." (Ludwig Wittgenstein, 1953, p. 223). Gill elucidates Wittgenstein, explaining that for Polanyi tacit knowledge is where the mind and body collapse into one, meaning we cannot speak Lion because "we do not share the embodied experiences, emotions, and cultural practices of the Lion's world" (Satinder P. Gill, 2015, p. 21). Nagel famously captured the essence of strong tacit knowledge in his essay *What Is It Like to Be a Bat?* writing,

> Reflection on what it is like to be a bat seems to lead us, therefore, to the conclusion that there are facts that do not consist in the truth of propositions expressible in a human language. We can be compelled to recognize the existence of such facts without being able to state or comprehend them.
>
> (Thomas Nagel, 1974, 441)

Alternatively, it could be asked *What Is It Like to Be a Human?* A machine cannot be human simply by behaviour, as it does not share the same embodied experience of flesh and blood, which Polanyi argued allows people to relate or empathise with each other and what he preferred to call "conviviality" (Michael Polanyi, 1961, p. 245). For example, strong AI can imitate the behaviour of pain (weak tacit knowledge), but humans personally know the experience of pain (strong tacit knowledge), which allows us to understand and relate to one another. Searle made this distinction with his famous Chinese Room Thought Experiment, focusing on the difference between speaking Chinese and knowing Chinese (John R. Searle, 1980, pp. 417–418).

Although Polanyi did not anticipate that machine learning would rely on weak tacit principles, he did appreciate that one day a machine "may conceivably simulate these propensities [think, feel, imagine...] to such an extent as to deceive us altogether" (Michael Polanyi, 1962, p. 263). However, Polanyi argued against the ontology implicit within the imitation game, explaining: "A deception, however compelling, does not qualify thereby as truth: no amount of subsequent experience can justify us in accepting as identical two things known from the start to be different in their nature" (Michael Polanyi, 1962, p. 263). When it comes to the nature of a machine, Polanyi makes an important distinction between a person and the things persons create, explaining:

> For a machine is a machine only for someone who relies on it (actually or hypothetically) for some purpose, that he believes to be attainable by what he considers to be the proper function of the machine: it is the instrument of a person who relies on it. This is the difference between machine and mind.
>
> (Michael Polanyi, 1962, p. 262)

Given that machines are created by persons, Polanyi argues that they work within a "tripartite system" as follows:

I	II	III
Mind→	Machine→	Things to which the machine informally refers.

(Michael Polanyi, 1952, p. 314)

Within (III), Polanyi explains that the machine informally refers to "functions, purpose etc., entertained by the mind," which he called operational principles. Thus, Polanyi concluded, a "machine can be said to function intelligently only by aid of unspecifiable personal coefficients supplied by the user's mind" (Michael Polanyi, 1962, p. 262).

There exists, according to Polanyi, an ontological dependency between persons and the machines they create that transpires from particulars into composites when setting their operational principles. This is demonstrated in that the particulars, or parts that make up the machine follow nature's laws, as studied by physics and chemistry, but the composites themselves, which make up the machine, operate according to laws imposed upon it to achieve a desired purpose, as studied within engineering. Polanyi argues that it is the operational principles that ontologically define a composite, such as a machine, via its purpose, explaining:

> The true knowledge of a machine which we have on the upper level is the understating of a *purpose* and of the *rational means* for achieving it; while the knowledge of its physical and chemical topography is by itself meaningless, for it lacks any conception of purpose or achievement. It becomes meaningful only when oriented towards establishing the material conditions for the success or failure of a machine.
>
> (Michael Polanyi, 2014, p. 52)

Intuitively, people do not define a created composite according to its parts but by its purpose. Polanyi uses the example of a watch to make this point, explaining:

> To understand a watch is to understand what it is for and how it works. The laws of inanimate nature are indifferent to this purpose. They cannot determine the working of a watch, any more than chemistry or physics of printers' ink can determine the concept of a book.
>
> (Michael Polanyi, 1969b, pp. 152–153)

Although a watch is composed of particulars that follow physical chemical laws, it is not and cannot be, defined by those parts. The reason for this is because the machine is a composite created by persons, and as such, follows the laws imposed upon those parts to accomplish its purpose, which in the case of a watch is keeping time. It is significant to note in Polanyi's argument the irreducible nature between particulars and composites; the purpose of a machine cannot be arrived at by studying the physical chemical laws of its parts. In the same way, a language cannot be learned by studying its letters alone. This goes back to Gödel's *incompleteness theorem,* in which the meaning of the composites, words, or mathematical symbols are defined outside the system by persons. According to Polanyi's ontology, an android is a composite and the intentional creation of its human creators, and as such, ought to be defined by its purpose. Thus, a machine that passes the Turing Test ought to be defined as a successful human imitation machine and not a human. Although a successful imitation machine is not human, it is capable of producing AD, which raises growing concern for a better understanding of what dignity is.

4. Dignity – Whose approach best supports it? Polanyi's or Turing's?

Polanyi and Turing have shown two opposing views. Turing argued for a human ontology of imitation that reduced to behaviour and Polanyi argued for a machine ontology that is not reductive but the product of the purposeful intention of humans. Thus far the UDHR's use of inherent dignity has managed to coexist with opposing philosophies, such as Polanyi's and Turing's, because it has remained, as explained by Luban "strategically silent about what key terms such as 'human dignity' are supposed to mean"; Luban subsequently concludes with the concern that "a concept that can mean anything means nothing" (David Luban, 2007, p. 68). This is particularly troubling given the UDHR's goal of stopping dehumanization and the mounting challenge to inherent human dignity from recent technological advancements. Kraynak argues that the threat of dehumanization from technology is a serious concern and argues that "the major challenge of our times is to recover a true and authentic understanding of human dignity and to defend it against threats from modern civilization" (Robert P. Kraynak, 2003, p. 2).

In light of Polanyi's development of tacit knowledge, it is not surprising that dignity continues to elude an explicit definition. This may be due to the tacit or non-reductive nature of human characteristics, such as intelligence, intentionality, and morality. Alternatively, clarity on what dignity is might be best gained through understanding whose worldview, Polanyi's or Turing's, best supports inherent dignity and what implications follow. It has already been demonstrated that Turing's approach erodes the *inherent* nature of human dignity; whereas, Polanyi upholds it by making a clear distinction between man and machine. Now it will be shown that Turing's approach also fails to support the *dignity* of humanity given the intrinsic moral duty owed I-Thou relationships as opposed to I-It relationships. Although Polanyi did not directly address human dignity, he did understand his approach as a possible foundation writing, "we see before us a way of knowing a human being in the fulness of his dignity through recognizing in him the same powers of understanding by which we are understanding him" (Michael Polanyi, 1961, p. 242).

Polanyi's entire argument shows that reductive physicalism fails to uphold human dignity by reducing humanity to either a random or determined collection of I-It parts; in reference to these two implications, he writes,

> It is simply this sort of mechanical reductionism that is the heart of the matter. It is this that is the origin of the whole system of scientific obscurantism under which we are suffering today. This is the cause of our corruption of the conception of man, reducing him either to an insentient automaton or to a bundle of appetites. This is why science denies us the possibility of acknowledging personal responsibility.
>
> (Michael Polanyi & Harry Prosch, 1975, p. 25).

Polanyi argued that persons are greater than their parts being relationally responsible to themselves and others. Humanity's responsibility to I-Thou relationships is a key aspect of dignity. Fundamentally, human dignity references moral obligations intrinsic to I-Thou relationships as opposed to I-It relationships. Polanyi's ontology is fundamentally committed to I-Thou relationships in which moral duty necessarily flows. That is, dignity references the moral obligation to value I-thou relationships so as to not reduce them into I-It relationships. The imitation game however is fundamentally based on this dehumanizing reduction of persons to I-It relationships, which neglects its moral duty to uphold I-Thou relationships. One reason for this is that a reductive ontology denies the existence of morality all together. There are however, philosophers, within the reductive camp, such as Wielenberg, who do argue for moral facts without reference to any persons, including God, but ultimately he fails to address what obligation one owes those facts.

This is the relational obligation that is inherent to dignity which Wielenberg acknowledges stating that, "Dehumanization is a kind of evil that strikes at the very heart of morality" (Erik J. Wielenberg, 2019, p. 137). Thus, he is implicitly arguing that it is morally wrong to treat an I-Thou relationship as an I-It relationship, but fails to explain what duty one owes these moral facts. The point being, duties are owed to persons (I-Thou) and not to moral facts (I-It). That is, I can owe my friend something, whereas I never owe my table anything. Taylor, a nontheist, makes this point writing, "A duty is something that is owed…. But something can be owed only to some person or persons. There can be no such thing as duty in isolation." (Richard Taylor, 1985, p. 83) Taylor understands the theistic implication of duty being grounded in a person, writing: "Our moral obligations can… be understood as those that are imposed by God…. But what if this higher-than-human lawgiver is no longer taken into account? Does the concept of a moral obligation… still make sense?…. the concept of moral obligation [is] unintelligible apart from the idea of God. The words remain but their meaning is gone" (Richard Taylor, 1985, p. 83-84). For this reason, Taylor gave up moral obligations and although Wielenberg can appeal to moral facts, there is no obligation to follow them.

Polanyi argued that his approach pointed the way to God, to which he found himself drawn, specifically to Christianity; he writes,

> I have mentioned divinity and the possibility of knowing God. These subjects lie outside my argument. But my conception of knowing opens the way to them. Knowing, as a dynamic force of comprehension, uncovers at each step a new hidden meaning. It reveals a universe of comprehensive entities which represent the meaning of their largely unspecifiable particulars. A universe constructed as an ascending hierarchy of meaning and excellence is very different from the picture of a chance collocation of atoms to which the examination of the universe by explicit modes of inference leads us. The vision of such a hierarchy inevitably sweeps on to envisage the meaning of the universe as a whole. Thus natural knowing expands continuously into knowledge of the supernatural.
>
> (Michael Polanyi, 1961, p. 246)

In order to account for one's moral obligation to I-Thou relationships (dignity) would metaphysically require a relational nature to moral facts that is best understood within a theistic framework. Specifically, it would require a personal lawgiver and those capable of understanding and obeying those relational laws. This places God into an I-Thou relationship with humanity and humanity into an I-Thou relationship with God, one's self, and others. Accordingly, dignity would be inherent to humanity due to God's I-Thou relationship with humanity and humanity's intrinsic capacity for I-Thou relationships as persons. Those relationships would be owed dignity given the nature of the moral law within an I-Thou relationship. Thus, it is morally wrong to treat an I-Thou relationship as an I-It relationship, which demonstrates that Turing's approach fails to support dignity and Polanyi's framework provides a possible foundation within Christian Theism.

Conclusion

The following paper has juxtaposed Polanyi and Turing to elucidate two machine ontologies, one based on purpose and the other based on imitation. The goal of this enquiry is to highlight dehumanizing concerns implicit to the imitation game. Granting humanness to machines represents a loss of the inherent nature of dignity, as specified by the UDHR, and opens the way for the dehumanization of certain humans. In order for humanity to retain inherent human dignity, anthropomorphized machines must *not* be viewed as human but as proposed by Polanyi 's ontology of purpose, a human imitation machine. Although a machine is not human, technological advancements in AI and robotics produce artificial dignity (AD), which create a need for greater clarity regarding what real dignity is. The UDHR leaves human dignity undefined but this might be necessary now and into the future, given the tacit nature of humanity. Alternatively, dignity can be elucidated through its metaphysical foundation, which is best supported through Polanyi's non-reductive approach of persons and founded in a moral lawgiver as demonstrated in Christian theism. If a dystopian future is to be avoided, human dignity must be upheld and a Polanyian ontology provides the necessary clarity to navigate further technological advancements in AI and robotics.

Works Cited

A. M. Turing. (1950). Computing Machinery and Intelligence. *Mind*, *49*, 433–460.

A. M. Turing. (1996). Intelligent Machinery, A Heretical Theory. *Philosophia Mathematica*, *4*(3), 256–260.

A Universal Declaration of Human Rights. (n.d.). Retrieved from https://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/217(III)

Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, & Sebastian Thrun. (2017). Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature*, *542*, 115–118.

Andrew Dobson jailed for "child-like" sex doll import bid. (2017). Retrieved from https://www.bbc.com/news/uk-england-stoke-staffordshire-40383627.

Andrew Hodges. (2009). Alan Turing and the Turing Test. In *Parsing the Turing Test: Philosophical and methodological Issues in the Quest for the Thinking Computer* (pp. 13–22). Springer.

CREEPER Act of 2017. , Pub. L. No. H.R. 4655.

David Luban. (2007). *Legal Ethics and Human Dignity*. New York: Cambridge University Press.

David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, … Demis Hassabis. (2017). Mastering the game of Go without human knowledge. *Nature*, *550*, 354–359.

Discussion on The Mind and the Computing Machine. (1949, October). Interdisciplinary discussion presented at the Manchester University. Retrieved from https://www.turing.org.uk/sources/wmays1.html

Erik J. Wielenberg. (2019). Evil and Atheistic Moral Realism. In *Explaining Evil: Four Views*. London: Bloomsbury Academic.

Ethem Alpaydin. (2016). *Machine Learning*. Cambridge, Massachusetts: MIT Press.

Jessica Riskin. (2016). *The Restless Clock: A History of the Centuries-Long Argument over What Makes Living Things Tick.* Chicago: The University of Chicago Press.

John Danaher. (2017). The Symbolic-Consequences Argument in the Sex Robot Debate. In *Robot Sex: Social and Ethical Implications* (pp. 103–131). Cambridge, Massachusetts: MIT Press.

John R. Searle. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, *3*(3), 417–457.

John R. Searle. (2009). The Turing Test: 55 Years Later. In *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer* (pp. 139–150). Springer.

Kevin Warwick, & Huma Shah. (2017). Taking the fifth amendment in Turing's imitation game. *Journal of Experimental & Theoretical Artificial Intelligence*, *29*(2), 287–297.

Ludwig Wittgenstein. (1953). *Philosophical Investigations*. Oxford: Basil Blackwell.

Michael C. Seto, & James M. Cantor. (2006). Child Pornography Offenses Are a Valid Diagnostic Indicator of Pedophilia. *Journal of Abnormal Pschology*, *115*(3), 610–615.

Michael Polanyi. (1952). The Hypothesis of Cybernetics. *British Journal for the Philosophy of Science*, *2*(8), 312–315.

Michael Polanyi. (1957). Scientific Outlook: Its Sickness and Cure. *Science*, *125*(3246), 480–484.

Michael Polanyi. (1961). Faith and Reason. *Journal of Religion*, *41*(4), 237–247.

Michael Polanyi. (1962). *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago: The University of Chicago Press.

Michael Polanyi. (1964). The Feelings of Machines. *Encounter*, (22), 85–86.

Michael Polanyi. (1966). *The Tacit Dimension*. Chicago: University of Chicago.

Michael Polanyi. (1969a). Knowing and Being. In Marjorie Grene (Ed.), *Knowing and Being: Essays by Michael Polanyi*. London and Chicago: Routledge & Kegan Paul ad University of Chicago.

Michael Polanyi. (1969b). The Logic of Tacit Inference. In Marjorie Grene (Ed.), *Knowing and Being: Essays by Michael Polanyi*. London and Chicago: Routledge & Kegan Paul ad University of Chicago.

Michael Polanyi. (2014). *The Study of Man*. Mansfield Centre, CT: Martino Publishing.

Michael Polanyi, & Harry Prosch. (1975). *Meaning*. Chicago: The University of Chicago Press.

Nan Li, Noboru Matsuda, William W. Cohen, & Kenneth R. Koedinger. (2015). Integrating Representation Learning and Skill Learning in a Human-like Intelligent Agent. *Artificial Intelligence*, *219*, 67–91.

Nick Haslam. (2006). Dehumanization: An Integrative Review. *Personality and Social Psychology Review*, *10*(3), 252–264.

Nick Haslam. (2014). What is Dehumanization? In *Humanness and Dehumanization* (pp. 34–48). New York: Psychology Press.

Nick Haslam, Brock Bastian, Simon Laham, & Stephen Loughnan. (2012). Humanness, Dehumanization, and Moral Psychology. In *The Social Psychology of Morality: Exploring the Causes of God and Evil*. Washington, DC: APA Press.

Peter Singer. (1975). *Animal Liberation* (second, 1990). New York: Avon Books.

Richard Taylor. (1985). *Ethics, Faith and Reason*. Englewood Cliffs, N.J.: Prentice-Hall.

Robert P. Kraynak. (2003). Defending Human Dignity: The Challenge of Our Times. In *In Defense of Human Dignity Essays of Our Times* (pp. 1–9). University of Notre Dame Press.

Satinder P. Gill. (2015). *Tacit Engagement: Beyond Interaction*. Springer International Publishing.

Sinziana Gutiu. (n.d.). *Sex Robots and Roboticization of Consent*. Presented at the We Robot Conference 2012. Retrieved from http://robots.law.miami.edu/wp-content/uploads/2012/01/Gutiu-Roboticization_of_Consent.pdf

Thomas Nagel. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, *83*(4), 435–450.

William Taussig Scott, & Martin X. Moleski. (2005). *Michael Polanyi: Scientist and Philosopher*. New York: Oxford University Press.