



BREAKING THE CODE: TOWARDS A CARE-BASED AI ETHICS

RICA DELOS REYES-ANCHETA

DEPARTMENT OF THEOLOGY
ATENEO DE MANILA UNIVERSITY
rancheta@ateneo.edu

AI Ethics is a nascent field that has emerged in response to growing concerns regarding the impact of artificial intelligence (AI) on individuals and societal structures. Digital ethics is part of the broad field of digital technology, which constantly evolves as thinkers explore the deployments of novel digital technologies. The rapid and dynamic nature of digital technologies requires serious ethical considerations, which involve an analysis of the Code of Ethics/ethical principles. This paper elucidates on AI ethics, drawing from Kantian theory, and bridges the gap by employing Care Ethics. It attempts to answer these questions: 1) What is AI ethics? 2) Using Kantian philosophy can AI Ethics refute the dignity-skeptics? 3) How can Care ethics bridge the gap in the AI Code of Ethics?

This paper argues that a principle-based AI Code of Ethics anchored in Kantian philosophy does not suffice to address ethical problems, and hence, extrapolates on a care-based AI ethics.

Keywords: Artificial Intelligence, Digital Ethics, Immanuel Kant, Care Ethics

Dates:

Received: October 17, 2024

Revised: March 4, 2025

Accepted: March 21, 2025

Published(Online): March 31, 2025

How to cite this article:

delos Reyes-Ancheta, R. (2025). Breaking the Code: Towards a Care-based AI Ethics. *Scientia - The International Journal on the Liberal Arts*, 14(1), 54–65. <https://doi.org/10.57106/scientia.v14i1.197>

Copyright:

Online: Asean Citation Index, DOAJ.
This work is licensed under the Creative Common Attribution License © 2025.
Print: Philippine Copyright © March 2025 San Beda University

Read online



Scan this QR code with your smart phone or mobile device to read online





“Sometimes the questions are complicated, and the answers are simple.” ~ Dr. Seuss

INTRODUCTION

AI Ethics is a nascent field that has emerged in response to growing concerns regarding the impact of artificial intelligence (AI) on individuals and societal structures. Digital ethics is part of the broad field of digital technology, which constantly evolves as thinkers explore the deployments of novel digital technologies. The rapid and dynamic nature of digital technologies requires serious ethical considerations, which involve an analysis of the Code of Ethics/ethical principles.

This paper elucidates AI ethics, drawing from Kantian theory, and bridges the gap by employing Care Ethics. It attempts to answer these questions: 1) What is AI ethics? 2) Using Kantian philosophy can AI Ethics refute the dignity-skeptics? 3) How can Care ethics bridge the AI Code of Ethics gap?

This article argues that a principle-based AI Code of Ethics anchored in Kantian philosophy does not suffice to address ethical problems, and hence, extrapolates on a care-based AI ethics.

ARTIFICIAL INTELLIGENCE AND ETHICS

Artificial Intelligence (AI) has gained popularity in recent years as ChatGPT was released in late 2022, the first easy-to-use generative artificial intelligence (GenAI) tool made widely available to the public.¹ It was followed by

more sophisticated versions fueling the race among large technology companies in the field of GenAI model development. Hence, many sectors are now paying more attention to digital technologies, like machine learning applications. However, AI poses multiple risks such as safety, data privacy, copyright, and manipulation of data. In view thereof, AI ethics in education poses questions about the ethical use of Generative AI. Granted that AI enables people to automate data than before, and technology develops faster than we thought a decade ago, the query of how we can responsibly use AI is often left to people’s devices.

The numerous AI ethics initiatives, documents, and guidelines over the past few years have primarily taken the shape of high-level abstract and prescriptive principles, set up by the European Commission, one of which is the “Recommendation of the Council on Artificial Intelligence” by the Organization for Economic Co-operation and Development (OECD).² More and more academicians want to understand the responsibility of users as AI continues to evolve and develop. As cited by Anna Jobin, Marcello Ienca, and Effy Vayena,³ academicians fear that AI might jeopardize jobs for human workers, be misused by malevolent actors, elude accountability, or inadvertently disseminate bias, and thereby undermine fairness. Several studies have discussed the topic of ethical and unintended negative consequences like algorithmic bias or discrimination. With the potential risks involved in using AI, who are liable, and can we handle the implications of AI? National and international organizations have responded to these societal fears by

¹ Tengrrl. 2024. “New: Proceedings of the Computers & Writing Annual Conference, 2023 (WAC Clearinghouse).” WPA-Announcements. February 3, 2024. <https://wpa-announcements.tracigardner.com/2024/02/03/new-proceedings-of-the-computers-writing-annual-conference-2023-wac-clearinghouse/>.

² Rodrigues, Rowena, Anais Resseguier, and Nicole Santiago. 2023. “When Artificial Intelligence Fails.” *Public Governance Administration and Finances Law Review* 8, no. 2 (December): 17–28. <https://doi.org/10.53116/pgaf.7030>.

³ Jobin, Anna, Marcello Ienca, and Effy Vayena. 2019. “The Global Landscape of AI Ethics Guidelines.” *Nature Machine Intelligence* 1, no. 9 (September): 389–99. <https://doi.org/10.1038/s42256-019-0088-2>.





developing ad hoc expert committees on AI, often commissioned with the drafting of policy documents. Universities followed suit and formed ad hoc committees to craft policies and guidelines to address the emerging ethical concerns in Generative AI tools. Interestingly, in Jobin, et al's study,⁴ they noted that there is no single ethical principle appeared to be common to the entire corpus of documents, although there is an emerging convergence around the principles of transparency, justice, fairness, non-maleficence, responsibility, and privacy. Similarly, UNESCO issued a document titled *Guidance for Generative AI in Education and Research*⁵ to support countries in implementing immediate actions, planning long-term policies, and developing human capacity to ensure a human-centered vision of these new technologies. As GenAI tools are increasingly able to automate some basic levels of writing and artwork creation, they are forcing education policymakers and institutions to revisit the ethical boundaries of the utilization of GenAI in myriad fields. These are critical considerations for education in this new phase of the digital era.

AI'S GENESIS & WRITING CODES

Coeckelbergh⁶ discovered that the term 'AI' was first introduced in 1955 by John McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. Generative AI refers to artificial intelligence that creates content automatically in response to prompts in natural-language conversational interfaces. It doesn't simply curate webpages or draw images to understand content, it can rather produce new

content that can appear in formats that comprise all symbolic representations of human thinking, which includes among others, photographs, digital paintings, cartoons, videos, music, and software code. GenAI is trained using data collected from various sources as long as they are available online. It generates its content by statistically analyzing the distribution of words, pixels, or other elements in the data. However, it cannot generate new ideas or solutions to real-world challenges, as it does not understand real-world objects or social relations that underpin language. Moreover, despite its fluent and impressive output, GenAI cannot be trusted to be accurate as it gathers information from a hodgepodge of data online.

During the early twenty-first century, numerous computer specialists and engineers aimed to create a strong AI system that mirrored human intelligence in every way except for its artificial nature. The feasibility of achieving this has sparked discussions for many years. The prominent American philosopher John Searle⁷ introduced the so-called Chinese room argument to contend that strong or general AI (AGI)—that is, building AI systems that could deal with many different and complex tasks that require human-like intelligence—is, in principle, impossible. In doing so, he sparked a long-standing general debate on the possibility of AGI. Current AI systems are narrowly focused (that is, weak AI) and can only solve one task, such as playing chess or the Chinese game of Go. Searle's general thesis was that no matter how complex and sophisticated a machine is, it will nonetheless have no 'consciousness' or 'mind', which is a prerequisite for the ability to understand, in contrast to the capability to compute (see section 2.e.). Searle's argument has

⁴ Ibid.

⁵ *Guidance for Generative AI in Education and Research*. 2023. UNESCO eBooks. <https://doi.org/10.54675/ewzm9535>.

⁶ Coeckelbergh, Mark. 2019. "Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability." *Science and Engineering Ethics* 26, no. 4 (October): 2051–68. <https://doi.org/10.1007/s11948-019-00146-8>.

⁷ Searle, John R. "On the Construction of a Social Reality." In *University of Iowa Press eBooks*, 63–64. <https://doi.org/10.2307/j.ctt20ks0rb.46.2018>.





been critically assessed against the counterclaims of functionalism and computationalism. It is generally argued that intelligence does not require a particular substratum, such as carbon-based beings, but that it will also evolve in silicon-based environments if the system is complex enough.⁸ Currently, some researchers have begun exploring AI that equates the concept with machine learning.

A hovering question on coding similarly emerged as AI developed. Can AI write codes or change the code snippets? Given the great potential of AI, we can surmise that AI-powered code editors use machine learning algorithms to analyze patterns in code. Bradford⁹ affirms the unimaginable potential of AI as he underscores that AI can predict and suggest code snippets for anyone, making the coding process faster and more efficient. It can break codes and generate new codes or entire functions based on user prompts, which can accelerate the development process. Similarly, Ruiz¹⁰ asserts that Generative AI coding tools are changing software for enterprises. She argued that it is evolving in ways we thought impossible—from code generation abilities to vulnerability detection and facilitating comprehension of unfamiliar codebases to streamlining documentation and pull request descriptions. In short, they are fundamentally reshaping how developers approach application infrastructure, deployment, and their work experience. An example to wit is GitHub Copilot assists developers in code completion, and with AI chatbots, AI code generation can be interactive and fun. However, the fact that code generation has become as simple as counting,

has people considered the ethical aspect of all these developments?

MORAL CONSIDERATIONS: MACHINES OR PEOPLE

There have been recent debates about AI ethics inquiring into who is ultimately accountable – machines or humans? Some have raised the important question of whether (a) current robots used in social situations or (b) artificial intelligent machines, once they are created, might have a moral status, and be entitled to moral rights as well, comparable to the moral status and rights of human beings.

The following discussion elaborates on the concept of moral status, freedom, and responsibility. The concept of moral status is vital to the development of ethics and moral philosophy because all beings who have a moral status are considered part of the moral community and are entitled to moral protection. Noting that all members of a moral community have moral status, they differ in the extent of knowledge and claims for moral protection. Additionally, a moral being has a moral status, which accords moral (and legal) rights as well. An example is how our perspective on the rights of ethnic minorities, women, and the LGBTQ+ community has changed over the years. We view the rights of animals and the environment differently now than a century ago. Hence, inclusivity is a highly valued quality and not just a moral demand. Given the development of artificially intelligent machines, we cannot rule out the possibility of attaching the notion of personhood to AI. This being, an ongoing debate, may be addressed as we shift our attention to the creator/s of AI.

⁸ Chalmers, David J. "Facing up to the Problem of Consciousness." In *The MIT Press eBooks*, 4–27. <https://doi.org/10.7551/mitpress/6860.003.0003>: 1996.

⁹ Bradford, L. *Using AI for Coding: Everything you need to know*. Learn to Code with Me. <https://learntocodewith.me/resources/ai-coding/> accessed February 20, 2024)

¹⁰ Ruiz, Jeimy. 2024. "How AI Code Generation Works - the GitHub Blog." *The GitHub Blog*, May 14, 2024. <https://github.blog/ai-and-ml/generative-ai/how-ai-code-generation-works/>.





Kant highlighted the notion of autonomy in the context of moral status and rights. Inherent in an autonomous being is the concept of dignity, which can be traced to Kant's seminal work, *Groundwork for the Metaphysics of Morals*, in 1785¹¹ where he argued that all persons have an inherent value, or dignity, in virtue of their rational autonomy. This value commands a distinct kind of moral respect, which we express by abiding by certain limits in our treatment of others. Thus, Kant argued that we have a categorical duty to treat persons always "as an end" and "never merely as a means." A moral person is defined as a rational and autonomous being. By this definition alone, it is possible to ascribe personhood to artificial intelligent machines once they have reached a certain level of autonomy, giving the machine the ability to make moral decisions. Current AI machines are becoming increasingly autonomous, so it seems only a matter of time until they meet this moral threshold. Why is this vital to understanding ethical stance? This begs the question: Is autonomy the sole criterion for moral status and rights?

Kant's view of "seeking out" the foundational principle helps identify the foundational element of moral status—the rational agent. A thinking being, a rational agent, can decide whether to act (or not act) by the demands of morality. The ability to make decisions may be feasible for a machine soon. However, determining what is good, regardless of environment or constraints, makes the person an autonomous agent. It is not the ability to make decisions that is a priori, but the being. Recognizing the person acting willfully and voluntarily ascribes the person to a moral status. Additionally, all rational agents have dignity insofar as they act autonomously. As a person acts with autonomy, s/he becomes morally responsible.

¹¹ Timmermann, Jens. *Kant's Groundwork of the Metaphysics of Morals*. <https://doi.org/10.1017/cbo9780511487316>: 2007.

Let us further elucidate on two relevant points. First, is to distinguish between "goodwill" and "duty." Wilburn¹² discussed that Kant considers "goodwill" the only intrinsically valuable thing. The will drives our actions and grounds the intention of our act. Hence, it is good in itself. Kant recognizes that it is difficult to determine one's intentions, and thus, he alludes goodwill to acting in conformity with duty and acting from duty. He also theorizes the natural behavior of people. He suggests that there are two types of moral duty. 1) categorical, and 2) hypothetical. Kant argues that one should act according to the law—the universal law. On the other hand, hypothetical imperative refers to the moral duty of a person as that person deals with personal interests, emotions, desires, and the like. There is a significant shift to consciousness here – where qualities of the person and emotions affect the desired action. Kant extrapolates on our rationality, which dictates that we act following our representations or conceptions of laws. However, what inspires us to formulate a maxim or a law is an end (4:427). This argument distinguishes a human being as opposed to an AI. There is an innate ability to regulate one's behavior following a maxim or principle.

Therefore, Kant underlines the foundational moral principle, which is rooted in each person's rational will. This means that, inherently, we are all bound by moral standards as thinking, autonomous beings. To assess one's actions and decisions and consider the inherent good makes the person unique and truly free. An autonomous person can distinguish right from wrong and leads the rational agent to seek the common good. This awareness of innate goodness and moral status attached to one's dignity humanizes a person. An important point

¹² Wilburn, H., PhD. *An introduction to Kant's moral theory*. Pressbooks. <https://open.library.okstate.edu/introphilosophy/chapter/a-brief-overview-of-kants-moral-theory/> accessed August 12, 2022).





here is that we share a common humanity, in which we know that “in your person as well as in the person of any other, always at the same time as an end, never merely as a means (4:429). Thus, AI serves humanity’s purposes in myriad ways. It cannot synergize with a full-blooded human being in the absence of physical attributes. After all, AI emerged to serve an end and not an end in itself. Kant puts it succinctly: So, act that you use humanity in your own person as well as in the person of any other, always at the same time as an end, never merely as a means.

Rigby¹³ raised ethical concerns about AI technology and underscored that it signals a crucial concern to balance the benefits and risks of AI technology. AI technology is integrated into the healthcare system, education, and industries. However, little has been explored about the ethical risks of AI implementation. Little too has been explored in the areas of privacy and confidentiality. How can human beings be protected from unwarranted algorithms?

‘CARING’ ETHICS, MORE THAN DIGNITY-DRIVEN ETHICS

Care ethics has long challenged the structures of domination and control. To recognize Care ethics as a viable ethical framework and practice requires that we situate AI in a long-entrenched culture of domination and power. Humans have raced for power and prestige since time immemorial. Common knowledge tells us that the powerful gain more influence on others and have greater access to resources as opposed to those who are categorized as “have-nots.” The powerless have fewer resources and limited access to finer things, much less to the vast array

of AI generative tools. Although Care ethics does not offer a quick antidote to the growing AI throes, it seeks the risks prevalent in the structural forms of control in AI. In advancing science and technology, people can devalue some people, like women and those who by religious, linguistic, racial, ethnic, and sexual minorities, categorize them as second-class citizens.

In this section, I argue that Kantian ethics is not enough to address the emerging AI issues and thus, propose Care ethics to be normative and up the ante in search of an applicable and context-sensitive ethical framework.

The fragility and complexities of life challenge us to consider altering our perspective of care ethics by adopting it as a normative and context-driven ethical framework.

A. MORAL NORMS

Moral norms preserve in propositional form the insights of the human experience of value, the experience of what helps or hinders the well-being of persons and all creation. Norms guide our moral choices, providing us with spaces for cognition and valuation from an array of choices before us. Hence, they may be positive or negative. The proscriptive and prescriptive natures of norms serve as indicators for our actions and decisions.

Moral theology informs us that there are two categories of norms, which relate to character or action. Formal (or general) norms relate to character. Material (or particular) norms, also called concrete behavioral norms, relate to action. While some norms are considered absolute, these same norms are also limited in a sense by factors such as cultures and historical epochs. Even if formal norms are intended to aid a person to be what he/she ought to be, norms also take into

¹³ Rigby, Michael J. “Ethical Dimensions of Using Artificial Intelligence in Health Care.” *The AMA Journal of Ethics* 21, no. 2 (February): E121-124. <https://doi.org/10.1001/amajethics.2019.121:2019>.





account the dynamism of the human person. Hence, norms are intended to exhort, challenge, and encourage people to develop or grow. A classic example of a norm is the golden rule—expressed on a more positive note in Mt.7:12;¹⁴ Lk.6:31¹⁵. Proscription fashions the language of norms in terms that prohibit or in terms of don'ts. The expressions of formal norms, however, may be more complex than they seem. For example, stealing affirms moral liability because the prescription not to steal has a moral demand on a person. However, the term carries a more complex meaning that must be seen along with other factors that affect that specific prescription. The terms that designate “stealing”, “killing” or “genocide” carry moral evaluation of that action, making the prescription: “do not steal” or “do not kill” more an injunction and categorizing these acts immoral. These acts are unjustifiable regardless of intention or circumstances, but they must be determined before a judgment can be made. Formal norms tend to absolutize the character and intention of an action.

Material norms relate to the sorts of actions we ought to perform. They attempt to attach formal norms to concrete pieces of behavior—to speech, to killing, to making promises, to sexual conduct, etc. Material norms, therefore, lead us to answer a practical moral question: “What must I do?” Addressing this question engenders further questions, which complicate a moral dilemma. For instance, a doctor has in some instances must hide the truth from a patient for health reasons. A patient who has a heart ailment may not be disposed to know further that a cancer cell exacerbates her illness. Hence, material norms are not yet moral judgments in the true sense. They would pronounce us guilty

of immorality when we do more harm (ontic evil/premoral disvalue) than necessary to achieve good.¹⁶ Material norms thus, make us cognizant that there are other factors constitutive of a moral obligation and thus, forewarn us not to make a quick judgment. Thus, a moral obligation takes us further than merely seeing the norms. It seeks to uncover the implicit and explicit norms.

B. IMPLICIT AND EXPLICIT NORMS

There are implicit norms, which may belong to the first-order and second-order classifications. The first-order norms are those do's and don'ts that people must obey in order to survive. These are the norms that belong to the call of nature or instincts. We thus normally say that we must pay heed to the call of hunger or thirst. If we do not eat, we die; if we do not drink water, we also die – so the norm is, “Take care of yourself.” These needs-calls are felt deep inside a person and by one's whole body. Other basic needs, borne out of other bodily needs, demand provisions for clothing or shelter as protection from the heat or cold. We may also say that making provisions for housing and clothing are obligations or “norms” for survival. Such “norms”, however, are first-order norms that need not be expressed and yet are to be “obeyed”. These “norms” could even be imposed on somebody or an institution capable of fulfilling the needs of the hungry or homeless. When this happens, the implied norm becomes explicit and necessary in settings where people have no previous relations and are not obligated to support or assist one another.

The second-order implicit norms belong to the call of society and culture to conform to standards. These are explicit norms to the uninitiated but implicit to their bearers or socially formed agents.

¹⁴ “Matthew 7:12 - Ask, Seek, Knock.” n.d. Bible Hub. <https://biblehub.com/matthew/7-12.htm>. “So, in everything, do to others what you would have them do to you, for this sums up the Law and the Prophets.”

¹⁵ “Luke 6:31 (NIV).” n.d. Bible Gateway. <https://www.biblegateway.com/passage/?search=luke%206:31&version=NIV>. “Do to others as you would have them do to you.”

¹⁶ Ibid., p. 74. See further discussions on norms by Fiona Robinson, *Globalizing Care Ethics: Feminist Theory and International Relations* (Colorado: Westview Press, 1999).





Parents or significant others explicitly teach the norms of society and culture to children who would later embody and externalize them. Once norms become second nature to people, there is no need to remind them of their obligation to conform. We can thus consider these norms turning implicit and become part of the behind-the-back standards. People thus behave according to the implicit norms that already form part of their dispositions. Thus, people's identities, qualities, pursuits, and hopes for the future are *also* formed and driven according to the implicit norms of society.

There are explicit norms that do not belong to the first-order or second-order norms but are generated by normative practices exemplified in the practices of some agents of society. Other explicit norms may be formulated by philosophers or academicians based on their logical interpretations of principles.

Commands given by parents or other significant authorities function as norms to children; prescriptions by somebody who is not considered the authority may only carry a normative force, but people are not compelled to obey it. People may be inclined to obey the prescriptions if it is clearly shown that refusal to follow will produce something bad or will deprive them of a good. When these reasons are established and validated by practice and conformity among a large population, the prescription may turn into a norm — explicit, initially, and could transform into something implicit in the future. Similarly, care ethics can transcend into a normative ethical framework if people recognize its inherent value. Let's consider this example: before germs were discovered to be responsible for many diseases, people did not have to have clean hands. Thus, the norm, "Wash your hands before eating," was alien to them. The awareness about germs has given existence to the norm. Another example is

lending money. To lend money with interest used to be considered wrong (it was usury) because it destroyed community solidarity and support systems based on compassion. In later years, due to the spread of the lending-with-interest practice in towns and cities, usury became loans and no longer proscribed. It is because of what is *perceived* as good or evil contained in practices and their effects on individuals and society, that practices carry with them the implied properties of norms. When such practices become validated by time, that is, when people see their utility, practicality, or "goodness" (fruits), norms become culturally or institutionally formalized.

Norms come in handy when needed to judge choices or acts at home, in the neighborhood, at work, in school, and in marketplaces. What is deemed wholesome and agreeable must somehow fit the standards, which are also translated into specific norms that categorically command or prohibit. The standards are somehow maintained firm and stable through the norms that are kept and observed. Thus, the commands: Do not do... Do... These come mostly as common-sensical prescriptions, implicit pre-established norms, or explicit rules and obligations.¹⁷

Individuals obey or disobey norms or taboos. The injunction "Do not kill," and "Do not lie," are some of the familiar norms which may be obeyed faithfully or disobeyed with impunity. Some cases of total disregard for norms do not, however, indicate that norms are not powerful. There are more average persons—those whom society would regard as civilized, normal, or socially well-mannered—than uncivilized or ill-mannered. Because of the normalized persons, we observe more instances of preserved social order than social disorder. Besides, those who disturb the social order through the breaking of

¹⁷ Dagmang, *The Predicaments of Intimacy and Solidarity*, pp. 269ff; see also pp. 29ff.





norms, as well as laws, are punished. Laws, in the punitive sense, ensure the observance of norms or compliance with rules.

A society, therefore, would need norms (with the corresponding laws) to maintain a social order. A classroom needs some rules or norms of decorum to maintain a certain environment conducive to learning. A society, in the same vein, would need some norms to promote desirable conditions for its continuity or proper functioning.

Persons who are properly raised in a certain society are those whose dispositions have been shaped to conform to popular norms and established laws. Without this conformity, individual integration into society will not be easy. That is why children are taught by their parents some acceptable forms of beliefs, actions, and rituals. It is also through the parents that society primarily continues to produce and reproduce itself and ensures the continuity of its traditions and history.

It is important, then, for carers to identify concrete measures, anticipate/forecast, design, analyze, and implement the actual caring needs in light of their peculiar contexts. In an interview with Harari,¹⁸ he said: “The really big revolution of the 21st century will not be in our tools, in our vehicles, in our society, or our economy. The really big revolution will be in ourselves, in humanity. The main products of the 21st-century economy will be bodies, brains, and minds. We are now learning how to hack not just computers but how to hack organisms, and in particular how to hack humans; we are learning how to engineer them and how to manufacture them.” Harari’s words echo what we all dread: fear of what AI can become and how it can revolutionize not only technology but also our value system.

¹⁸ Harari, Y. N. *Homo Deus: A Brief History of Tomorrow*. (New York: Harper, 2017).

Thus, care ethics is more relevant given this scenario as it attempts to unhinge us from the ontological view of the human being as autonomous and transform our social lens that dehumanizes people to see clearly that we are inherently all ‘connected’ with others. This sense of ‘connectedness’ is the defining feature of the human condition. In universalistic morality, we rely on deductive reasoning and the brilliance of arguments to merit moral value to a sound position. As AI in education calls for more than the logical process to sensitize people to use AI for good. It must seek a deep understanding of the value of the human person and hence, AI must serve its purpose of being a ‘tool,’ and not replace humans or mimic people’s behavior. Crafted by a thinking being, homo sapiens, the crafted tool must gain a collective “moral status” attached to its maker. Thus, for example, if AI is utilized to gather and manipulate information available online into unverified data sources that are prompted by a human person, it defies its ethical responsibility and goes against the value of knowledge production.

Care ethics challenges thinkers and those who develop AI Generative tools to embed collective moral parameters that can guide and enhance the ability of these tools. They are not ultimately to determine if such tools are to be penalized or sanctioned. To question the way AI is used is to acknowledge that care places accountability on technology. It is admitting that we all have an inherent duty to care for the welfare of the present and future generations, not merely as dignified beings, but as homo sapiens. Our innate ability to determine our flaws and limitations transcends our crude behavior and makes us wise agents. Thus, if we don’t take cognizance of the dangers of data collection and algorithms, who would?





Care veers away from deontological morality. It affirms that the care deficit we experience in a fast-paced World Wide Web is a problem we must deal with, beginning with how we assess AI tools. Technological advancement does not emerge for no reason. As thinking beings, we grasp the purpose of any invention or innovation. If we allow it to go unchecked, care ethics will slowly rescind human consciousness and autonomy. To critique and analyze is a form of caring.

Fisher and Tronto¹⁹ identified the four dimensions of care, namely, caring about (recognizing caring needs in the first place) and care receiving (being responsive to the ways); caring with (thinking about the effects of multiple care processes on trust and respect) and care always occurs in a context. They argued that some simple and familiar examples of needs are defined by more powerful members of society rather than those who have needs themselves. The latter informs us that the development of AI stems from the more powerful people's needs and not necessarily the majority. Big firms need to speed up the processes and mechanisms to gain more profit. Big tech companies want to double their income to keep the economic advantage and up the ante in the name of competition. The more advanced technologies are, the more reputable and successful companies become. No wonder, AI is becoming more sophisticated and can become more autonomous, just as fictional novels portray. When do we take seriously the moral compass to be 'careful?' When do we also measure up to the moral demand to care?

Valuing care has a social dimension. It urges us to care with and grow in caring. Caring is an activity that is less valued or deemed parochial.

It rarely gets beyond the caring abode or caring circle. People see its weak positioning against deontological, consequential, or virtue ethics. But what must drive caring? How can it be an ethical stance?

In simple terms, caring is altruism. When we care about the people closest to us, it isn't a virtue at all. It simply is caring for people we care about—a default reaction to people who love us. It becomes 'care ethics' when it springs from a sense of otherness. When one is not compelled to do the right thing but chooses it out of care for the other who is unknown to you. That is caring for others as a fundamental response to a fellow human being's need. This requires a consciousness of the other's rights and needs, which makes the act virtuous. It doesn't necessarily require a response, and yet, a person chooses to respond to attend to another person's needs. Hence, this archetypal response is a caring response that makes humanity flourish. Imagine a world where people are drawn to caring, imbued with a desire to go beyond the ethical call of justice.

Sevenhuijsen²⁰ pointed out that care must be valued as a social good. This perspective is crucial to any reform as she pushes caring to the public domain. This would change the common belief that care ethics is parochial and relegated to home ethics alone. I believe that we all want a free society where people are free to care and where caring relationships exist. The moral imperative to care can redirect the emerging need for AI development. As AI's potential reaches the roof, people must come from a place of caring to understand the future possibilities and throes of data collection and analysis.

¹⁹ Tronto, J. C., & Fisher, B. Toward a Feminist Theory of Caring. In E. Abel, & M. Nelson (Eds.), *Circles of Care* (pp. 36-54). (SUNY Press:1990).

²⁰ "Citizenship and the Ethics of Care: Feminist Considerations on Justice, Morality, and Politics." n.d. Routledge & CRC Press. <https://www.routledge.com/Citizenship-and-the-Ethics-of-Care-Feminist-Considerations-on-Justice-Morality-and-Politics/Sevenhuijsen/p/book/9780415170826>.





It is no secret that blanket surveillance is becoming common. In addition, much of the data is traded between agents or agencies for a fee. Data collection, sale, and use are paid for a sum of money or to perpetuate the manipulation of data. Harris²¹ identified the main data collection for the ‘big 5’ companies, such as Amazon, Google/Alphabet, Microsoft, Apple, and Facebook). He also opined that most of these companies are based on deception as they fight for a ‘surveillance economy’. It is no longer surreal to generate a person’s replica using face recognition in photos and videos or do profiling through algorithms. More and more data shall yield information about society, politics, daily life, and people. Harari²² further suggested a shift in control—as he describes a data-centric world where computers, robots, and artificial intelligence (AI) reign supreme. This bleak forecast speaks of the long-term implications of AI. He inquires: “What will happen to society, politics, and daily life when non-conscious but highly intelligent algorithms know us better than we know ourselves?”

What happens when deep fake text becomes more credible than the truth? Unless caring becomes a mandate for the citizenry, it would be impossible to rely on cognitive-driven digital ethics. Policies on privacy and manipulation will struggle to catch up with technical and technological advancements. State and business actors will continue to watch and perhaps, manipulate people with the aid of AI technology, advancing their business interests instead of yielding the truth. We can also look to a future where opacity and bias are central issues as predictive analytics replace discernment and objectivity. Algorithms can dictate our daily affairs following the patterns or data trails we

leave online. AI systems will soon generate plausible actions and dictate decisions over time, altering even our preference for a worker—where robots can get the job done better.

CONCLUSION

AI ethics is a vast field that leaves black holes for ethicists to explore. Kantian’s turn to autonomy is not the one-size-fits-all moral philosophy that can solve the moral dilemma brought about by AI’s development. It is caring *for* others that opens the possibility for mutual concern and collective accountability. While AI Generative tools have made people’s lives easier and more efficient, they also uncover blind spots that appear in the chasm of power-intoxicated agents. Hence, harping on the centrality of moral duty is inadequate without Care ethics. Caring can reveal the contexts of the AI tools in question and re-focus on the creator of AI, which is the *homo sapiens*, not the *homo deus*.

Adopting care ethics can set a moral limit to the utilization of AI tools and further exploration of AI’s potential. The inadequacy of rules and policies can be addressed by caring human agents who choose to acknowledge the need for a deep consciousness and accountability based on the inherent capacity of humans to care.

²¹ Harris, T. How technology is hijacking your mind—from a magician and Google design ethicist. *medium.com*, Thrive Global (18.05.2016).

²² Harari, Y. N. *Homo Deus: A Brief History of Tomorrow*. (New York: Harper, 2017).





REFERENCES:

- Banner, W. A. (1981). *Moral Norms and Moral Order: The Philosophy of Human Affairs* (p. 1). University Press of Florida.
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, Opportunities and Challenges toward Responsible AI. *Information Fusion*, 58(1), 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Bradford, L. (2024, February 20). Using AI for Coding: Everything You Need To Know. Learn to Code with Me. <https://learntocodewith.me/resources/ai-coding/>
- Coeckelbergh, M. (2019). Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability. *Science and Engineering Ethics*, 26(4). <https://doi.org/10.1007/s11948-019-00146-8>
- Cole, D. (2020). The Chinese Room Argument (Stanford Encyclopedia of Philosophy). Stanford.edu. <https://plato.stanford.edu/entries/chinese-room/>
- Gula, R. M. (1982). What are They Saying about Moral Norms? (p. 54). Paulist Press.
- Harari, Y. N. (2016). *Homo Deus: A Brief History of Tomorrow*. Harper Perennial.
- Harris, T. (2016). How technology is hijacking your mind—from a magician and Google design ethicist. Medium.com; Thrive Global . <https://medium.com>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56–67. <https://doi.org/10.1038/s42256-019-0138-9>
- Resseguier, A., & Rodrigues, R. (2021). Ethics as attention to context: recommendations for the ethics of artificial intelligence. *Open Research Europe*, 1, 27. <https://doi.org/10.12688/openreseurope.13260.1>
- Rigby, M. J. (2019). Ethical Dimensions of Using Artificial Intelligence in Health Care. *AMA Journal of Ethics*, 21(2), E121–124. <https://doi.org/10.1001/amajethics.2019.121>
- Robinson, F. (1999). *Globalizing Care : Ethics, Feminist Theory, and International Relations*. Westview Press.
- The Proceedings of the Annual Computers and Writing Conference, 2023. (2024). The WAC Clearinghouse EBooks. <https://doi.org/10.37514/pcw-b.2024.2296>
- Tronto, J. C., & Fisher, B. Toward a Feminist Theory of Caring. In E. Abel, & M. Nelson (Eds.). (1990). *Circles of Care* (pp. 36–54). SUNY Press.
- Wilburn, H. (2022, May 30). An introduction to Kant's moral theory. Open.library.okstate.edu; Pressbooks. <https://open.library.okstate.edu/introphilosophy/chapter/a-brief-overview-of-kants-moral-theory/>

