

A Sartrean Perspective On Human Freedom And Moral Responsibility In The Age Of Artificial Intelligence

Ediana-abasi Louis Bassey

Department of Philosophy, University of Uyo, Nigeria

Article received: 30/05/2026, Article Accepted: 15/06/2026, Article Published: 24/06/2026

DOI: - <https://doi.org/10.55640/ijsshd-v03i06-06>

© 2026 Authors retain the copyright of their manuscripts, and all Open Access articles are disseminated under the terms of the [Creative Commons Attribution License 4.0 \(CC-BY\)](#), which licenses unrestricted use, distribution, and reproduction in any medium, provided that the original work is appropriately cited.

ABSTRACT

The rapid advancement of artificial intelligence (AI) has generated significant philosophical and ethical concerns regarding human freedom, autonomy, and moral responsibility. As AI systems participation in decision-making processes across healthcare, education, transportation, governance, and other sectors, raises questions concerning accountability for actions and outcomes influenced by intelligent technologies. This study investigated the concepts of human freedom and moral responsibility in the age of artificial intelligence from the perspective of Jean-Paul Sartre's existential philosophy. The problem addressed is whether the growing influence of AI threatens human freedom and whether responsibility for AI-assisted decisions can be transferred from human agents to technological systems. The major aim of the study was to examine the distinction between human and machine decision-making, evaluating whether AI can possess freedom in the Sartrean sense. Others are to analyze the issue of responsibility for AI-generated outcomes, and assess the relevance of Sartre's concept of bad faith in contemporary digital environments. The study is justified by the increasing integration of AI into critical areas of human life and the need for a philosophical framework capable of addressing emerging questions of agency and accountability. Its significance lies in contributing to ongoing discussions in AI ethics, existential philosophy, and technology governance. The research adopted a qualitative philosophical methodology based on conceptual and critical analysis of primary and secondary literature. The scope of the study is limited to the examination of freedom, responsibility, authenticity, and bad faith within the context of AI-assisted decision-making. The study concluded that artificial intelligence lacks the consciousness, intentionality, and existential freedom necessary for moral agency.

KEYWORDS

Artificial Intelligence, Jean-Paul Sartre, Freedom, Moral Responsibility, Existentialism, Bad Faith, Human Agency, AI Ethics.

INTRODUCTION

Imagine a hospital that uses artificial intelligence to assist doctors in diagnosing patients. One day, the system recommends a treatment plan that ultimately results in a patient's death. An investigation is carried out and a difficult question emerges; who is responsible? Is it the doctor who approved the recommendation, the programmers who designed the system, the hospital that adopted the technology or the artificial intelligence itself? As artificial intelligence becomes involved in human decision making, questions such as these have become more frequent and more complex. Similar questions can also arise from a case of an accident occurring in a self-

driven car programmed with Artificial Intelligence. Should the owner of the self-driven car be held responsible or the programmers or the artificial intelligence itself?

The integration of artificial intelligence into human decision-making has generated uncertainty concerning the nature of freedom, agency, and accountability. As AI systems become capable of making recommendations and performing tasks previously reserved for human beings, there is growing confusion regarding who should be held responsible when harmful outcomes occur. Some scholars and practitioners increasingly attribute decision-

making authority to AI systems, raising concerns about the erosion of human autonomy and moral responsibility. Despite extensive discussions in AI ethics, there remains insufficient philosophical engagement with these issues from an existentialist perspective. This study therefore investigates whether AI challenges human freedom and whether moral responsibility can legitimately be transferred from human beings to intelligent machines.

The rapid advancement of Artificial Intelligence (AI) has transformed various aspects of human life, including education, healthcare, communication, and governance. AI systems are increasingly capable of performing tasks that traditionally required human intelligence, thereby raising important philosophical and ethical questions concerning human freedom, autonomy, and moral responsibility. As intelligent technologies become more integrated into everyday decision-making processes, scholars have begun to question whether human beings remain fully autonomous agents in an increasingly automated world. One of the central concerns in contemporary AI discourse is the extent to which human beings can be held responsible for actions and decisions that are influenced by intelligent systems. While AI technologies can assist in decision-making, they also create challenges regarding accountability, agency, and ethical responsibility. These concerns become particularly significant when AI systems produce harmful outcomes or influence human choices in subtle but powerful ways.

Sartre argues that human beings are fundamentally free and that every individual is responsible for the choices he or she makes (Sartre 29). According to his famous principle that “existence precedes essence,” human beings are not born with a predetermined nature but continuously create themselves through their actions (Sartre 22). Consequently, individuals cannot evade responsibility by blaming external circumstances, since doing so would constitute what Sartre describes as “bad faith” (Sartre 47). Although AI influences human choices and behavior, Sartre’s philosophy maintains that human beings remain fundamentally free and therefore morally responsible for the decisions they make regarding the creation, use, and regulation of AI technologies.

This study examines human freedom and moral responsibility in the age of artificial intelligence from a Sartrean perspective. It seeks to determine whether the increasing use of AI threatens human freedom and whether individuals, institutions, and societies remain morally accountable for decisions involving AI technologies.

Existence Precedes Essence

Jean-Paul Sartre’s philosophy of value is mostly in his works *Being and Nothingness* (1943) and *Existentialism Is a Humanism* (1946). These works present a radical

rethinking of morality in a post-theistic world. Sartre rejects the traditional grounding of value in divine command, natural teleology, or rational necessity and instead locates value within human freedom itself. Value, for Sartre, is neither discovered nor objectively embedded in the structure of reality; it is constituted through conscious choice. At the heart of Sartre’s philosophy lies the claim that “existence precedes essence” (Sartre, 28). This thesis, central to Existentialism Is a Humanism, declares that human beings first exist and only later define themselves through their actions. Sartre defines essence as “The sum of the formulae and qualities which made its production and its definition possible” (Sartre 29). While the idea of essence coming first is possible when it has to do with say furniture, where a chair can be made with a particular purpose, when it comes to human beings, there is no predetermined human nature, no divine blueprint, and no intrinsic moral law governing conduct. Instead, everyone creates his or her own essence through lived experiences, decisions taken, choices made, therefore existence comes before essence.

To Sartre, when a person is born, he or she is simply “thrown” into the world. overtime, identity is formed, character is developed, and the person gradually discovers his or her own essence through choices made regularly. As he puts it, “Man is nothing else but what he makes of himself” (28) this shows that man has a choice. Therefore, if truly existence is prior to essence man is therefore responsible for what he is. According to Ezedike (220), “To say that existence precedes essence is a radical affirmation of human freedom.”

Freedom

The idea of thrownness into the world for human beings to form their own identity and purpose through choices leads to a significant implication of the idea of Freedom. This means that since there is no predetermined essence, human beings are at liberty to choose their path in life. In *Being and Nothingness* Sartre famously argues that consciousness is characterized by its ability to transcend given circumstances and project itself towards future possibilities unlike objects which simply are what it is. He consequently maintains that human beings are “condemned to be free” because having been thrown into existence without his consent, he remains responsible for everything he does. He cannot avoid choosing, in fact not choosing is already seen as a choice. Ultimately, freedom is an unavoidable condition of a human being. This freedom comes with responsibilities.

Freedom does not entirely mean that man is free from all constraint because they may be faced by what Sartre calls “facticity” which refers to the condition to which a person is born into such as race, social class, religion, historical influence however, these conditions do not ultimately determine human action. Individuals reserve the freedom to transcend these situations through their choices.

(Being and Nothingness, 511) He further argues that there is no external authority, divine command, and fixed human nature that commands nor determine what an individual ought to be. According to Sartre, "Man is nothing else but that which he makes of himself" (Existentialism is a Humanism, 29). Scholars such as Thomas Flynn have emphasized the importance of freedom in Sartre's philosophy, saying that freedom is not one characteristic among many but rather the defining feature of human existence making responsibility unavoidable. (Flynn 53), similarly David Detmer argues that Sartre's conception of freedom requires individuals to recognize themselves as active agent rather than passive products of circumstance (Detmer, 78). These observations reemphasize Sartre's belief that man must assume responsibility.

Responsibility

Freedom and responsibility are inseparable. Since human beings create themselves through choices, they must also accept responsibility for the consequences of the choices they make. responsibility is one of the most fundamental implications of Sartre's philosophy of value. In *Existentialism Is a Humanism*, Sartre argues that when an individual chooses a course of action, he or she is not only choosing for himself, but he is also affirming a particular vision of what humanity should be. Every choice expresses a value judgment and serves as an example of human conduct, this is like Immanuel Kant's categorical imperative which asks if the principle of one's action can be used as a universal law. Consequently, Sartre says "In choosing for himself he chooses for all men." (Sartre 29) "To choose between this or that is at the same time to affirm the value of that which is chosen; for we are unable ever to choose the worse. What we choose is always the better; and nothing can be better for us unless it is better for all." (Sartre, 29). When one chooses, the choice speaks of the value one places on what was chosen. Consequently, by the virtue of choice, Sartre maintains the individuals are responsible not only for themselves but for humanity at large. This responsibility contributes to shaping the world morally. Man, however often attempt to avoid such responsibilities and try to push the blame on other factors such as artificial intelligence in the case of this discussion. This blame game is known as bad faith.

Bad Faith

Bad faith represents man's attempt to escape the implications of existence, freedom, choice, and responsibility. For Sartre, human beings are fundamentally free beings who continuously create themselves through their choices. However, the awareness of this freedom often produces anxiety and anguish, leading individuals to deny or conceal their freedom to avoid responsibility. It is this self-deception that Sartre calls bad faith. The roots of bad faith can be

traced to Sartre's famous principle that existence precedes essence. Because existence precedes essence, human life is characterized by choice.

The inevitability of choice gives rise to freedom. However, freedom inevitably leads to responsibility. Sartre insists that people are responsible not only for themselves but also for the values they affirm through their actions (Sartre, *Existentialism Is a Humanism* 29). Every decision reflects a judgment about what is valuable, desirable, or worthwhile. Thus, responsibility becomes the unavoidable counterpart of freedom. It is at this point that bad faith emerges. Rather than accepting responsibility for choices, the individual adopts excuses that portray actions as determined by external factors such as social roles, emotions, institutions, or circumstances (Sartre, *Being and Nothingness* 86). They try to say things like, "I had no choice, the system made me do it, the AI made the decision" In this way, bad faith is an effort to escape the anxiety associated with self-determination.

Sartre illustrates this phenomenon through several examples. One of the most famous is that of a waiter who performs his role so mechanically and completely that he treats himself as nothing more than a waiter rather than a free human being capable of choosing otherwise (Sartre, *Being and Nothingness* 102). Thomas Flynn argues that bad faith represents an attempt to reduce human beings to objects governed by external forces rather than acknowledging their status as free subjects (Flynn 61). Similarly, Jonathan Webber notes that bad faith involves a distortion of one's self-understanding in order to avoid the demands of freedom and responsibility (Webber 74). These interpretations highlight Sartre's belief that self-deception is not merely a psychological error but an ethical failure to acknowledge one's freedom. Having explored different concepts that determine one to be seen as a human being with regard to Sartre's existentialism, the question is does AI possess these qualities that can ultimately qualify it to be responsible for consequences of action it performs?

Artificial intelligence

According to Boucher, artificial intelligence is an umbrella term for computational systems that perform tasks associated with human intelligence such as learning, perception, decision-making, and language processing through rule-based programming, statistical learning, or biologically inspired optimization methods, rather than through human consciousness or understanding (Boucher 1–2).

Artificial Intelligence (AI) functions through the interaction of data, algorithms, computational power, and learning mechanisms that enable machines to perform tasks typically associated with human intelligence. At its core, AI seeks to replicate aspects of human cognition,

including learning, reasoning, perception, problem-solving, and decision-making. Russell and Norvig describe AI as the study of intelligent agent systems that perceive their environment and act in ways that maximize their chances of achieving their objectives (7).

The first stage in the operation of AI is data acquisition. AI systems require large amounts of information from which they can learn patterns and relationships. These data may include text, images, sounds, videos, numerical records, or sensor readings. The quality and quantity of data significantly influence the effectiveness of an AI system because machine intelligence depends on the information available to it for learning and adaptation. After data collection, AI systems process and organize the information through algorithms. Algorithms are sets of instructions that guide the computer in identifying patterns, correlations, and trends within data. Unlike traditional computer programs that follow fixed instructions, many AI systems employ machine learning techniques that enable them to improve their performance through experience. Machine learning allows a system to adjust its internal parameters based on new information, thereby enhancing its ability to make accurate predictions and decisions over time (Russell and Norvig 651).

One of the most significant developments in AI is the emergence of neural networks and deep learning. Neural networks are computational models inspired by the structure of the human brain. They consist of interconnected nodes, often called artificial neurons, which process information collectively. Deep learning systems contain multiple layers of neurons that enable them to recognize highly complex patterns in data. Through repeated exposure to examples, these systems learn to identify features, classify information, and generate outputs with increasing accuracy. Such systems are responsible for many modern AI applications, including facial recognition, language translation, speech recognition, and autonomous vehicles.

Another crucial aspect of AI is decision-making. Intelligent systems evaluate available information and select actions that best achieve predetermined goals. According to Russell and Norvig, a rational agent acts in a manner that is expected to maximize its performance measure based on the evidence available to it (35). In this sense, AI systems do not merely store information; they use information to make choices. For example, a medical diagnostic AI analyzes patient data and determines the most probable diagnosis, while a navigation system evaluates traffic conditions to recommend the fastest route. AI also relies on feedback mechanisms. During training, the system compares its outputs with desired outcomes and adjusts its parameters accordingly. This process, known as optimization, allows the AI to reduce errors and improve future performance. Continuous learning enables some AI systems to adapt to changing environments and new information, making them

increasingly effective over time.

From a philosophical perspective, the functioning of AI raises important questions about the nature of intelligence itself. Philosophers have long debated whether intelligence consists merely of information processing or whether it requires consciousness, understanding, and subjective experience. John Searle argues that computers manipulate symbols according to formal rules without genuinely understanding their meaning, suggesting that AI may simulate intelligence rather than possess it (417). His famous “Chinese Room” argument challenges the claim that computational processes alone can produce genuine understanding. In contrast, Alan Turing proposed that if a machine’s responses are indistinguishable from those of a human being, it may be reasonable to attribute intelligence to the machine (442).

This perspective focuses on observable behavior rather than internal consciousness. The debate between these positions remains central to contemporary philosophy of artificial intelligence. The operation of AI therefore involves more than technological processes. It intersects with questions concerning knowledge, reasoning, consciousness, ethics, and human identity. While AI systems can analyze information, learn patterns, and make decisions with remarkable efficiency, philosophers continue to question whether these capabilities constitute genuine intelligence or merely sophisticated simulations of intelligent behavior. Such questions are as follows:

- i. If AI systems can make decisions based on data and algorithms, what distinguishes human decision-making from machine decision-making?
- ii. Can a machine that processes information and produces intelligent outputs be said to possess freedom.
- iii. If humans increasingly rely on AI for decisions, are they surrendering part of their freedom and responsibility?
- iv. Who should be held morally responsible when an AI system causes harm the machine, the programmer, the organization, or the user?
- v. Can responsibility be transferred to an artificial system, or does moral accountability always remain with human beings?
- vi. Does the growing influence of AI encourage the tendency to avoid responsibility by blaming external forces for our actions?
- vii. If AI can imitate human reasoning and creativity, does this challenge Sartre’s claim that human beings are unique because they consciously create their own essence through choices?

- viii. Can an AI system ever possess consciousness, self-awareness, or intentionality, which Sartre considers essential to human existence?
- ix. How can human beings remain authentic and autonomous in a world increasingly shaped by algorithmic recommendations and automated decisions?
- x. In an age of artificial intelligence, what does it mean to be truly free and morally responsible?

Artificial Intelligence and Human Freedom

The explanation of artificial intelligence reveals that AI systems can learn from data, recognize patterns, and make increasingly sophisticated decisions. However, these capabilities raise important philosophical questions about freedom, agency, and accountability. While AI may simulate intelligent behavior, it does not possess consciousness or moral awareness. This distinction becomes significant when examined through Jean-Paul Sartre's existentialist philosophy, which emphasizes human freedom, personal choice, and responsibility. The emergence of AI therefore invites a reconsideration of whether human beings remain fully accountable for their actions in a technologically mediated world or whether they risk escaping responsibility by attributing decisions to intelligent machines.

Human Decision-Making Versus Machine Decision-Making

A common question in recent debates about artificial intelligence is whether machine decision-making can be compared to human decision-making. As AI systems become increasingly capable of performing tasks that resemble human reasoning, some argue that machines are beginning to make decisions in ways similar to humans. However, from a Sartrean perspective, there is a fundamental difference between the two. While humans make decisions as free and self-conscious beings, machines operate according to programmed instructions, algorithms, and data patterns. Therefore, what appears to be machine "decision-making" is fundamentally different from human choice.

Sartre's philosophy begins with the claim that human beings are conscious beings whose existence precedes their essence (Existentialism is a Humanism 22). Because humans are not defined by a predetermined nature, they must continually choose who they will become. Every decision is therefore an act of self-creation. When a person chooses a career, speaks the truth, tells a lie, or refuses to act, that choice contributes to the formation of his or her identity. Human decision-making is thus existential and self-defining. Machine decision-making, by contrast, lacks this existential dimension. An AI system does not choose its goals, values, or purposes.

Another important distinction concerns consciousness. In *Being and Nothingness*, Sartre argues that human consciousness is characterized by intentionality which is the ability to direct itself toward objects, possibilities, and future projects (21). Human beings are aware of themselves as agents making choices. They can reflect upon their actions, question their motives, and imagine alternative possibilities. This reflective awareness enables individuals to evaluate their decisions morally and existentially. AI systems do not possess such consciousness. They process information but do not experience the meaning of that information. An AI system may identify patterns in medical data and recommend a treatment, but it does not understand illness, suffering, or the ethical implications of its recommendation. It has no subjective awareness of the decision it produces. Therefore, while human decisions arise from conscious deliberation, machine outputs emerge from computational processes devoid of self-awareness.

The issue of responsibility further distinguishes human and machine decision-making. Sartre argues that because human beings are free, they are responsible for their choices and the consequences that follow from them (Sartre, *Existentialism Is a Humanism* 29). Individuals can be praised, blamed, rewarded, or punished because they act as conscious agents capable of choosing otherwise. Responsibility presupposes freedom and awareness. Machines cannot satisfy these conditions. An AI system cannot be held morally accountable because it lacks consciousness, intention, and freedom. It cannot experience guilt, remorse, obligation, or moral conflict. As Shannon Vallor observes, technological systems do not possess moral agency in the same sense as human beings because they cannot understand or assume ethical responsibility for their actions (Vallor 118). Consequently, responsibility for AI-generated outcomes remains with the human actors who design, deploy, supervise, and use these systems.

Sartre's philosophy provides a clear framework for distinguishing human decision-making from machine decision-making. Human beings possess consciousness, freedom, intentionality, and responsibility, all of which make genuine choice possible. Machines, regardless of their sophistication, operate through computational processes and lack the existential characteristics necessary for authentic decision-making. As a result, the moral responsibility associated with AI systems remains fundamentally human rather than technological.

The Question of Freedom

From the perspective of Jean-Paul Sartre, artificial intelligence does not possess freedom in the genuine sense of the term. For Sartre, freedom is not simply the ability to select between alternatives rather, freedom is an ontological condition of human existence. Human beings

are free because they are conscious beings capable of transcending their present circumstances and projecting themselves toward future possibilities (Sartre, *Being and Nothingness* 567). Human consciousness, which Sartre calls the for-itself (*pour-soi*), is never fixed or complete. It is constantly engaged in self-creation through choices and actions. This capacity for self-transcendence is what makes genuine freedom possible. AI lacks the essential characteristics that Sartre considers necessary for freedom. Sartre's distinction between being-in-itself (*en-soi*) and being-for-itself (*pour-soi*) further clarifies this issue. Being-in-itself refers to objects that simply exist without consciousness or self-awareness. They are what they are and cannot transcend themselves. Being-for-itself, on the other hand, refers to conscious beings who are aware of themselves and capable of self-determination (Sartre, *Being and Nothingness* 98). Human beings belong to the latter category because they can reflect upon their existence and redefine themselves through their choices.

Consequently, since machines lack the prerequisite for freedom, they therefore do not take responsibility for the consequence of AI actions. The absence of freedom also means the absence of moral responsibility. Sartre maintains that freedom and responsibility are inseparable because individuals are accountable for the choices they make (Sartre, *Existentialism Is a Humanism* 24). Since AI does not freely choose its actions, it cannot be held morally responsible for them. Responsibility remains with the human agents who create, deploy, regulate, and use AI systems. Shannon Vallor similarly argues that technological systems do not possess the moral agency required for ethical accountability, even when their actions have significant consequences (Vallor 112). For Sartre, freedom is not merely acting without external control; it is consciously choosing oneself through one's actions. Unless AI develops authentic self-awareness and existential subjectivity, it cannot possess freedom in the Sartrean sense.

The Question of Responsibility

Another pressing ethical questions in the age of artificial intelligence is the issue of responsibility when AI systems make mistakes. As AI increasingly participates in decision-making processes in healthcare, finance, education, transportation, security, and governance, harmful outcomes occasionally occur. An AI system may deny a qualified applicant employment, misdiagnose a patient, recommend a harmful action, or generate biased decisions. In such situations, the question arises: Who should be held responsible? From a Sartre's existentialism, the answer is clear: responsibility remains fundamentally human because AI lacks the freedom, consciousness, and moral agency necessary for accountability.

Sartre's philosophy is built upon the principle that human

beings are radically free and therefore responsible for their actions. In *Existentialism Is a Humanism*, he argues that every individual is responsible for the choices he or she makes because there is no external force that can completely determine human action (29). Freedom and responsibility are inseparable. Since human beings freely choose their actions, they must also accept the consequences that follow from those choices.

Contemporary AI ethicists support this conclusion. Shannon Vallor argues that technological systems do not possess moral agency and therefore cannot assume ethical responsibility for their actions (Vallor, 118). Similarly, Luciano Floridi contends that accountability for AI outcomes must remain with the human agents and institutions involved in the design and governance of technological systems (Floridi 142). These perspectives align closely with Sartre's insistence that responsibility belongs to free and conscious beings rather than to tools or machines.

However, Sartre's philosophy does not imply that responsibility always rests with a single individual. AI systems often involve complex networks of actors, including programmers, engineers, corporate leaders, regulators, and end-users. Responsibility may therefore be distributed across multiple human agents. What remains constant is Sartre's insistence that responsibility cannot be transferred to the machine itself because the machine lacks the existential conditions necessary for moral agency.

Sartrean analysis firmly rejects the notion that artificial intelligence can be morally responsible for its mistakes. Responsibility requires freedom, consciousness, and the capacity for self-determination, all of which AI lacks. When AI systems cause harm, accountability rests with the human beings who design, deploy, supervise, and use them. Attempts to attribute responsibility solely to AI constitute a form of bad faith because they obscure the human freedom underlying technological decision-making. Thus, in the age of artificial intelligence, Sartre's philosophy serves as a powerful reminder that technological advancement does not diminish human responsibility; rather, it increases the need for ethical vigilance and accountability.

Human Authenticity and Autonomy

The question of human authenticity and autonomy in an algorithm-driven world cannot be answered by saying that human freedom has been eliminated. Instead, from Sartre perspective this paper argues that freedom is always present, but it can be concealed, weakened, or denied through bad faith. Algorithmic recommendation systems therefore do not remove human autonomy; rather, they create new conditions in which humans are tempted to abandon responsibility for their choices.

Sartre's philosophy begins with the claim that human beings are radically free. In *Being and Nothingness*, he argues that consciousness is always self-transcending and never fully determined by external forces, meaning that human beings are always responsible for what they do with their situation (Sartre, 567). Similarly, in *Existentialism Is a Humanism*, Sartre insists that individuals are not defined by a fixed nature but must continually create themselves through their choices (Sartre, 22). From this standpoint, authenticity is achieved when individuals fully recognize and accept this freedom rather than denying it.

However, Sartre also acknowledges that human beings exist within facticity one of which is being born in an era of digital technologies. Algorithmic recommendation systems belong to this sphere of facticity. They influence what people watch, read, buy, and believe by filtering information and presenting curated options. In this sense, algorithms shape the "situation" in which choices are made, but they do not eliminate the ability to choose itself (Sartre, *Existentialism Is a Humanism* 29). A critical concern is not influence but self-deception. The convenience of automated recommendations may reduce critical reflection, leading individuals to outsource judgment to technological systems. This can result in a subtle form of self-deception in which people believe their preferences are fully shaped by algorithms rather than recognizing their ongoing responsibility for accepting or rejecting those suggestions.

Authenticity, in Sartrean terms, requires the opposite attitude. An authentic individual recognizes that although algorithms influence options, they do not determine choice. The person remains responsible for whether to follow, ignore, or critically evaluate recommendations. Authenticity therefore involves resisting the temptation to treat algorithmic systems as authorities that relieve one of responsibility. It means acknowledging: "I am still the one choosing, even within a digitally structured environment."

Scholars of technology ethics support this concern. Shannon Vallor argues that digital systems can shape habits of attention and decision-making in ways that weaken moral reflection, but they do not eliminate human agency (Vallor 118). Luciano Floridi similarly emphasizes that AI systems operate within "onlife" environments where human responsibility is continuously exercised rather than replaced by technology (Floridi 142). These views align with Sartre's insistence that technological structures modify conditions but never abolish freedom.

Sartre perspective, algorithmic recommendation systems do not eliminate human authenticity or autonomy, but they intensify the risk of bad faith. Human beings remain free, but in a digitally mediated world, freedom must be actively recognized and exercised. Authenticity today

requires resisting passive consumption and reaffirming personal responsibility in the face of technological influence. The challenge is not the absence of freedom, but the temptation to deny it.

The Risk of Bad Faith

Bad Faith (*mauvaise foi*) occurs when individuals deceive themselves by denying their own freedom and responsibility, often by attributing their actions to external forces. In the age of artificial intelligence, this concept becomes especially relevant because AI systems increasingly shape human decisions in subtle and powerful ways. However, based on Sartre's existentialism, it is arguable that while AI influences choices, it does not remove freedom; instead, it creates new opportunities for individuals to fall into bad faith. Sartre's philosophy is grounded in the idea that human beings are radically free. In *Being and Nothingness*, he explains that consciousness is not determined like objects in the world but is always capable of transcending its situation and choosing differently (Sartre, 567). Because humans are free, they are also fully responsible for their actions. Bad faith emerges when individuals attempt to escape this responsibility by denying their freedom.

In contemporary digital environments, AI systems influence human behavior through recommendation algorithms, predictive analytics, personalized content, and automated decision-making. These systems shape what people see, consume, and engage with. For example, social media platforms recommend videos, news, and advertisements based on user behavior. While these systems do not force decisions, they strongly guide attention and preference. For instance, constantly viewing contents on music leads to more content on music being brought to your timeline. What a social media user consistently see is what he spends time viewing previously. The length of time used in viewing and the frequency of viewing a certain content contributes to how often same or similar content will appear on your feed.

The danger arises when individuals begin to interpret this influence as deterministic control. Instead of recognizing that they are still choosing, people may say: "The algorithm made me do it., I watched it because it was recommended, I bought it because the AI suggested it." These statements reflect a denial of freedom. Sartre would classify this as bad faith because the individual is attributing personal choices to an external system in order to avoid responsibility (Sartre, *Being and Nothingness* 86).

Even though AI systems influence options, they do not eliminate the ability to reflect, question, or reject those options. one is not likely to see a government penalize AI for content that appears on a user's feed irrespective of the fact that he or she may have generated it using

artificial intelligence. Sartre's concept of facticity is important here because these conditions shape the context of decision-making but do not determine choice. As Sartre argues, individuals are always responsible for what they make of what is made of them (Sartre, *Existentialism Is a Humanism* 29). Thus, even in algorithmically structured environments, humans remain accountable for their responses.

Shannon Vallor argues that digital systems can shape habits of attention that weaken moral agency if users fail to remain critically engaged (Vallor 118). Luciano Floridi also notes that AI systems operate within human-designed frameworks, meaning accountability always traces back to human decision-makers (Floridi 142). These views support the claim that technology influences but does not replace human responsibility. A key insight is that bad faith is not caused by technology itself, but by how humans interpret and respond to it. AI becomes a convenient justification for avoiding the anxiety of freedom. Sartre calls this anxiety "anguish" the realization that we are fully responsible for our choices without external guarantees (Sartre, *Existentialism Is a Humanism* 25). Algorithmic systems can make it easier to avoid this anguish by giving the illusion that decisions are externally determined.

The influence of AI leads to bad faith not by removing freedom, but by encouraging individuals to deny it. When people attribute their choices to algorithms, they risk treating themselves as passive objects rather than free agents. Sartre's philosophy therefore serves as a warning: in an AI-driven world, the greatest ethical danger is not technological control, but human self-deception. Authentic living requires continuously recognizing that, despite algorithmic influence, individuals remain responsible for their choices.

Conclusion

This study examines the question of human freedom and moral responsibility in the age of artificial intelligence through the lens of Jean-Paul Sartre's existential philosophy. The research demonstrated that despite the growing sophistication of AI systems, artificial intelligence does not possess the essential characteristics that Sartre associates with authentic freedom, namely consciousness, self-awareness, intentionality, and the capacity for self-determination. While AI can process information, learn from data, and generate decisions that appear intelligent, it remains fundamentally different from human beings who continually create themselves through conscious choices.

Drawing from Sartre's principles of existence preceding essence, radical freedom, responsibility, and bad faith, the study argued that moral accountability for AI-generated outcomes remains with human beings rather than technological systems. Programmers, designers,

institutions, policymakers, and users cannot legitimately transfer responsibility to AI because such an act constitutes a form of bad faith a denial of human freedom and accountability. The increasing reliance on AI therefore does not eliminate human responsibility; rather, it heightens the need for ethical vigilance and conscious decision-making. Furthermore, it establishes that algorithmic systems influence human choices without determining them. Human beings remain free to accept, reject, or critically evaluate AI-generated recommendations. Consequently, authenticity in the digital age requires individuals to recognize their continuing freedom and resist the temptation to regard themselves as passive products of technological systems.

Sartre's existentialism offers a valuable philosophical framework for understanding contemporary debates surrounding AI ethics, agency, and accountability. As artificial intelligence becomes more integrated into social, economic, and political life, the preservation of human freedom and moral responsibility remains essential. Technological advancement should therefore be guided by ethical principles that affirm human agency rather than diminish it. The future of AI governance must continue to recognize that while machines may assist human decision-making, responsibility for those decisions remains fundamentally and irreducibly human.

References

1. Boucher, Philip. *How Artificial Intelligence Works*. European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA), European Parliament, Mar. 2019.
2. Detmer, David. *Freedom as a Value: A Critique of the Ethical Theory of Jean-Paul Sartre*. Open Court, 1988.
3. Floridi, Luciano. *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford UP, 2023.
4. Ezedike, Edward Uzoma. "Sartre's phenomenological ontology and existential ethics: A Critique." *Sapientia Journal of Philosophy*. Vol. 10, no 20, 2019.
5. Flynn, Thomas R.: *Existentialism: A very short introduction*. Oxford up, 2006.
6. IBM. *What is artificial intelligence (AI)?* IBM, www.ibm.com/topics/artificial-intelligence. Retrieved June 23, 2026.
7. Russell, Stuart J., and Peter Norvig. *Artificial Intelligence: A Modern Approach*, 4th ed., 1980, pp. 417-24.

- 8.** Sartre, Jean-Paul. Being and Nothingness. An Essay on Phenomenological Ontology. Translated by Hazel E. Barnes, Washington Square Press, 1992.
- 9.** Sartre, Jean-Paul. Existentialism Is a Humanism. Translated by Carol Macomber, Yale UP, 2007.
- 10.** Searle, John R. "Minds, Brains, and programs." Behavioural and brain sciences, vol. 3, no. 3, 1980, pp. 417-24.
- 11.** Turing, Alan M. "Computing Machinery and Intelligence." Mind, vol 59, no. 236, 1950, pp. 433-60.
- 12.** Vallor, Shannon. Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting. Oxford UP, 2016.
- 13.** Webber, Jonathan. The Existentialism of Jean-Paul Sartre. Routledge, 2009.