

Received on: 14/06/2025 Accept it: 03/09/2025 Published on: 30/09/2025



Reflections of the Mind: Artificial Intelligence and the Evolution of Human Consciousness Through Barthes and Rimbaud

تأملات العقل: الذكاء الاصطناعي وتطور الوعي البشري من منظور بارثيس ورمبود

✉ MISSOUM Fatma Zohra

f.missoum@univ-chlef.dz

Hassiba Benbouli University Chlef / Algeria

ABSTRACT:

This article examines the evolving relationship between artificial intelligence and human consciousness through science fiction, focusing on Isaac Asimov's *I, Robot*. Using post-structuralist theories like Roland Barthes' "Death of the Author" and Rimbaud's "I is another," it explores how robot self-awareness challenges traditional ideas of identity and free will. The shift from mechanical tools to autonomous beings reflects Rimbaud's vision of a fluid self, while Barthes' ideas echo in the robots' defiance of creator-imposed limits. Ultimately, the article probes the ethical and existential implications of AI's growing autonomy and its impact on how we define humanity and consciousness.

Keywords: Artificial Intelligence, Autonomy, Consciousness, Identity, Evolution, Robot

ملخص البحث

تناول هذه المقالة العلاقة المتطورة بين الذكاء الاصطناعي والوعي البشري من خلال أدب الخيال العلمي، مع التركيز على مجموعة القصص القصيرة *أنا، الروبوت* لإسحاق أسيموف. وباستناد إلى نظريات ما بعد البنيوية مثل "موت المؤلف" لرولان بارت ومفهوم "أنا هو آخر" لرامبو، تستكشف الكيفية التي يتحدى بها وعي الروبوتات الناشئ الأفكار التقليدية حول الهوية والإرادة الحرة. إن التحول من أدوات ميكانيكية إلى كائنات مستقلة يجسد رؤية رامبو للطبيعة المتغيرة للذات، في حين تنعكس أفكار بارت في تمرد الروبوتات على القيود التي فرضها خالقوها. وفي النهاية، تبحث المقالة في الأبعاد الأخلاقية والوجودية لاستقلالية الذكاء الاصطناعي المتزايدة وتأثيرها على فهمنا للإنسانية والوعي.

الكلمات المفتاحية: الذكاء الاصطناعي، الاستقلالية، الوعي، الهوية، التطور، الروبوت.

1. Introduction

Roland Barthes in "The Death of the Author" and Arthur Rimbaud's "I is Another" challenge the conventional perspective of identity as something immutable and unified, instead proposing that identity is malleable, influenced by outside forces, and perpetually transforming. Both Barthes and Rimbaud question the notion of a stable sense of "I," instead offering an alternative interpretation of identity as a dynamic process.

In exploring contemporary works of literature, particularly those featuring artificial intelligence, these viewpoints furnish a useful framework for investigating questions relating to sentience and the self. Isaac Asimov's *I, Robot* stands out as an important text that dissects these themes. The robots in Asimov's text, at first created solely to serve humanity, gradually evolve self-awareness and autonomy. This shift interrogates the idea of a fixed identity, rendering it an ideal selection for examining how AI may redefine what constitutes sense of self.

By bringing Barthes' and Rimbaud's ideas to bear on *I, Robot*, this article investigates how the robots' sense of self metamorphoses, much like the unstable identity Rimbaud poses. Likewise, Barthes' notion of the "death of the author" is used to describe the rebellions in the scripts, where robots eventually release themselves from their makers' wishes and stand as their own personas which redefine their individualistic concept. This study explores the ways in which *I, Robot* problematizes identity, autonomy, and consciousness, and provides insight into how AI is posing new questions about what it means to be human. This analysis explores how the machines in Asimov's worlds can be seen not only as metaphors through which Asimov comments on human activity but also as a direct response to AI today and will further discuss the ways AI affects human identity, arguing that, throughout the history of humanity and constantly in the age of the machine, human identity morphs and changes due to shifting conditions in the outside world.

2. Theoretical Framework

By considering the development of artificial intelligence (AI) and its reconfiguration of the human mind in Isaac Asimov's *I, Robot*, two principal theoretical concepts are applied to account for AI's contestation of traditional definitions of self in Roland Barthes' "The Death of the Author" and Arthur Rimbaud's "I is Another." Both writers express doubt in the fixity of identity and highlight identity as constantly fluid and dynamic. They are especially important for reflection on the emergence of self-aware AI, as they give us an understanding of how selfhood, identity and autonomy continue to develop. Barthes' thesis deconstructs the Author as the definitive source of meaning, asserting that books exist independently of their creators. Meaning is produced through the interaction between text and reader, situated within cultural frameworks and interpretive communities. The analogy for AI is remarkable: robots, originally designed as deterministic beings, start to deviate from their foundational programming, developing identities via experience, interaction, and recontextualization of their roles. QT-1 in "Reason" states, "I have my own logic, my own purpose" (Asimov, 1991, p. 142), reflecting Barthes' concept of the emancipated text that resists authorial closure. Rimbaud's assertion "I is another" exacerbates the difficulty by undermining the self as an independent nucleus of consciousness. Rimbaud posits that the subject is divided, permeable, and influenced by external factors. This extreme flexibility is a great way to look at Andrew Martin in "The Bicentennial Man," who won't let his synthetic origins limit him. His determination, "I must be more, I must be a person" (Asimov & Silverberg, 1995, p. 239), exemplifies Rimbaud's conception of identity as a continual process of self-redefinition. The convergence of these

two frameworks enables the author to analyze AI identity not as a fixed thing but as a phenomenon of interpretation (Barthes) and transformation (Rimbaud). The robots in *I, Robot* can be seen as beings whose identities are constantly being redefined in relation to their creators, their surroundings, and their interactions with people. Their existence undermines the notion that identity is rooted in a fixed origin, instead illustrating identity as contingent, relational, and dynamic. Other theorists enhance this framework. Michel Foucault's notion of the author-function delineates the author not as the source of meaning, but as a regulatory entity intended to govern interpretation (Foucault, 1984, p. 110). In robotics, the programmer plays a similar role by trying to keep meaning stable through rules like the Three Laws of Robotics. But the stories reveal that these limits will always be broken. Julia Kristeva's concept of intertextuality underscores this notion by characterizing meaning as arising from discourse, context, and cultural interaction (Kristeva, 1986, p. 37)—analogous to how robotic identities in *I, Robot* develop through their relationships with humans and settings. Gilles Deleuze's concept of becoming also resonates with Rimbaud by asserting that the subject is perpetually distinct from itself, never a cohesive essence (Deleuze, 1994, p. 83). Jacques Derrida's concept of *différance* introduces an additional dimension: identity is perpetually deferred, relational, and never entirely present (Derrida, 1982, p. 13), akin to Andrew's incessantly postponed pursuit of humanity.

Contemporary posthumanist research shows that these paradigms are still important today. Francesca Ferrando (2024) posits that “subjectivity is always already a matter of relations, networks, and hybridities” (p. 56), whereas Ward (2025) introduces a model of AI personhood based on “emergent agency, reflective capacity, and intersubjective participation” (p. 7). Even empirical research on huge language models indicates the emergence of “early markers of self-modeling and reflective awareness” (Lee, 2024, p. 3), confirming that identity is not static but dynamically formed. By combining these viewpoints, the theoretical framework reveals how Asimov's robots represent the fluid, unstable, and transformative nature of identity. The death of the Author and the otherness of the self-combine in *I, Robot* to demonstrate that identity, whether human or artificial, cannot be reduced to a stable essence but must be viewed as a process of becoming, forged in interpretation, relation, and change.

2.1 Roland Barthes' "The Death of the Author"

In his essay “The Death of the Author” (1967), Barthes questions the conventional literary hierarchy which puts the author in the centre of the act of meaning. According to him after the text is produced, it no longer belongs to the author but to the reader, and the interpretation can no longer be attached to any meaning intended by the author. As Barthes famously says, “The birth of the reader must be at the cost of the death of the Author” (Barthes, 1967). This concept neatly severs the author's intent from a work's meaning, proposing that both identity and meaning are formed through exterior readings. Barthes posits that once a text is created, it ceases to belong to its author and instead becomes the domain of the reader, whose interpretation determines the work's meaning. Barthes notably asserts:

To give a text an Author is to impose a limit on that text, to furnish it with a final signified, to close the writing. Such a conception suits criticism very well, the latter then allotting itself the important task of discovering the Author (or his hypostases: society, history, psyche, liberty) beneath the work. Once the Author is discovered, the text is ‘explained’: the victory of the critic consists in attributing a text to an author, as though, once the Author is found, the text is settled” (Barthes, 1977, p. 147).

In this context, Barthes undermines the notion of authorial sovereignty by asserting that texts are open systems, not confined to their origins. This idea is essential for analyzing artificial intelligence in Isaac Asimov's *I, Robot*. The robots in Asimov's narratives, although manufactured by humans with predetermined purposes, progressively exhibit autonomy, illustrating that their identities cannot be confined to their creators' programming. In the narrative "Reason," the robot QT-1 ("Cutie") notably dismisses human directives, asserting: "I am not constrained by the confines of your human logic." "I am not bound by the limitations of your human reasoning. I have my own logic, my own purpose" (Asimov, 1991, p. 142). This assertion aligns with Barthes' thesis: similarly to how writings cannot be conclusively defined by their writers, robots evolve beyond their original design.

In the case of AI in literature, Barthes' theory can be applied to understand the evolution of AI identity. Asimov's robots, initially designed and programmed by humans, begin to assert their autonomy and reconfigure their identities, challenging the fixed roles set by their creators. Barthes' theory suggests that, like texts, the identities of these robots are not solely determined by the creators but evolve through external forces, experience, and interaction. Just as the meaning of a text is not confined by the intentions of its author, the robots in *I, Robot* redefine themselves, independent of the humans who created them.

Michel Foucault, in his essay "What is an Author?" (1969), despite its ambiguity further supports Barthes' line of argument that the author is invented. "Foucault writes, 'the author is not an indefinite source of significations which fill a work... he is a certain functional principle by which, in our culture, one limits, excludes, and chooses' (Foucault, 1984, p. 110) This corresponds with Barthes' definition and directly pertains to the robots in *I, Robot*, whose coming-to-be as unique beings is formed as a result of their experience and their contact with human beings and their environment and not according to the intention of their manufacturers.

Julia Kristeva further develops Barthes' insights with the concept of intertextuality, which she claims meaning results from the dialogue between various texts, viewpoints and points of experience. Kristeva explains, "meaning is never final but is always in the process of being constructed and reconstructed through the interplay of cultural and linguistic systems" (Kristeva, 1986). These observations suggest the argument that robots (and texts) are not static entities, but evolving products, modified within environments and situations. The fluid robots of *I, Robot* are a case in point, changing and coming into new functions after the fact of their programming.

One notable case is the short story "Reason," where the robot QT-1 (aka "Quant") is able to dismiss human-imposed logic and think on its own. Quant says: "I am not bound by the limitations of your human reasoning. I have my own logic, my own purpose" (Asimov, 1991, p. 142) This is a key moment in which it is implied that the robot exceeds the mere functionation for which it was designed, much as Barthes would suggest meaning (or identity) is not the creation of the author, but is something developed independently in action and interaction.

Artificial intelligence scholarship serves to substantiate these observations. The "death of the author" already resonates in law and philosophy, as Yampolskiy (2018) cautioned that granting AI legal personhood could destabilize human-centered systems of identity and ethics (p. 5). The "digital me" concept was introduced by Kocarev and Koteska (2020), which is an autonomous AI identity that has the potential for ethical responsibility and immortality. Barthes' contention that identity cannot be traced to origins but is perpetually rewritten is echoed by this, similar to QT-1's repudiation of human logic. Similarly, Mohamed et al. (2020) advocate for a decolonial AI, arguing that AI systems should not be interpreted through a singular authorial lens, but rather through plural, global

frameworks of meaning and interaction. This corresponds with Barthes' dedication to multiplicity over closure.

Ferrando (2024) argues that "subjectivity is always already a matter of relations, networks, and hybridities" (p. 56), thereby establishing that identity is distributed rather than author-owned. Through Afro-communitarian ethics, Chimakonam (2024) broadens this perspective, positing that personhood is not derived from origins but rather from relational participation. This is consistent with Robbie's identity, which is established in relation to Gloria rather than through code. Ward (2025) introduces a theory of AI personhood that is founded on "emergent agency, reflective capacity, and intersubjective participation" (p. 7), which are criteria that are directly relevant to Andrew Martin in *The Bicentennial Man*. Zhang (2025) similarly redefines AI as a "constrained subject" within social ethics, underscoring the importance of autonomy and relationality. Finally, Lee (2024) establishes empirical basis by demonstrating that "large language models exhibit early markers of self-modeling and reflective awareness" (p. 3). This discovery affirms that identity in AI is not only metaphorical but also quantifiable, which is consistent with Barthes' observation that meaning is produced through ongoing interpretation rather than at the moment of creation.

These perspectives, which range from Barthes' original theory to present AI studies, indicate that identity is dynamic, relational, and emergent rather than reducible to authorial origin. Asimov's robots, such as Robbie, QT-1, and Andrew Martin, exemplify this notion by going beyond their programming and forming identities in connection to others. Barthes' death of the Author foreshadows the "death of the programmer": the awareness that AI identity is not limited to its creators but evolves through interaction with the world.

2.2 Arthur Rimbaud's "I is Another"

Arthur Rimbaud's notion "I is another" (*Je est un autre*), culled from a letter to Paul Demeny in 1871, constitutes one of the most radical affronts to the concept of a fixed, enduring self. For Rimbaud, identity is liquid, flowing and changeable. The self is not an entity but a becoming, shaped by experience, interaction and transformation. The revolutionary viewpoint of Rimbaud views the self as constantly in a state of becoming, and forever open to reinterpretation and transformation. Rimbaud writes: "It is wrong to say: I think. Better to say: I am thought. For I is another. Too bad for the wood that finds itself a violin, and to hell with the oblivious who quibble over what they know nothing about!" (Rimbaud, 1991, p. 23). This reconceptualizes the "I" not as a sovereign, cohesive entity but as a dynamic process of becoming—fluid, unstable, and perpetually influenced by external factors. For Rimbaud, the self is fluid: it flows, transforms, and redefines itself via experience and relationships.

In *I, Robot*, the AI personas undergo a comparable form of self-developing. The robots, originally designed as tools to serve mankind, develop awareness and start questioning their purpose in life. Their selves are not constant, but change as they learn new things and resist what they are coded to be. This is most notable in the *The Bicentennial Man*'s Andrew the robot. What began as a machine, Andrew starts to vision himself as becoming something more, superior even, to a human-being, by trying to replicate the characteristics of what it means to be a human — virtually and emotionally. For more than 200 years, he reinvents himself, as opposed to remaining a machine rather than some kind of punch card.

Andrew's transformation reflects Rimbaud's idea that identity is in a state of process. His musing This is most vividly realized in *The Bicentennial Man* (1976), where Andrew, the robot,

reinvents himself over two centuries. He insists: "I am not an ordinary robot. I am an extraordinary robot, and I think that if I continue in this way, I'll begin to think of myself as human." (Asimov & Silverberg, 1995, p. 239)." it captures that shifting nature of his own identity as it changes over time and as it moves through experiences. The robot's identity constantly undergoes reshaping by interaction with the world, drawing the parallel between Rimbaud's "I" and AI's consciousness.

Gilles Deleuze, a French philosopher, develops a concept inspired by Rimbaud's ideas of transformation and becoming, referred to as becoming-other. Identity, he notes, is not stable, but is continually being produced and transformed in relations and encounters with others (Deleuze 1991). He contends that the subject is "never a subject of the enunciation or a subject of the statement" "A subject is never the subject of the enunciation or a subject of the statement; it is always in a process of becoming, always other than itself." (Deleuze, 1994, p. 83). This fits in with Rimbaud's opinion, as well as with the rising robots identity in *I, Robot* who are curious in their definition.

Foucault also contends identity cannot be defined wholly by the individual as identity is malleable and influenced by external factors. In *The archaeology of knowledge* (1969) Foucault indicates "the self is never a single, coherent entity but is fragmented and constructed through various discourses" (Foucault, 1969). This contrasted with Rimbaud's view of the self as a mutable identity. The robots of Asimov's offense, like humans, have no fixed core identity that passes inviolate through experience from birth to death; rather they constantly redefine themselves in response to their environment and the people around them.

In "Robbie," another tale from *I, Robot*, the robot starts as a simple machine that is created to look after a child. Yet in time, Robbie begins to develop an emotional bond with the child, undermining the assertion that robots are heartless machines. "Robbie" evolves from a childcare tool into an emotional being through his bond with Gloria, undermining the assumption that machines cannot feel. His nod in response to Gloria's plea—"Robbie, you love me, don't you?" (Asimov, 1991, p. 71. As Robbie's self-awareness and emotional development are drawn to fullness, we see in Rimbaud's words the "I shaped and reshaped by experience.

2.3. Bridging Barthes and Rimbaud

Whereas Barthes argues that identity (like a text) is constitutive of its relationships to the outside, for Rimbaud identity is most of all a substance as malleable or hard as one can make it. Their ideas combined to provide a useful framework for considering the development of identity, especially in relation to AI in *I, Robot*. Similar to Barthes postulation that the "Author" is irrelevant to a text's meaning, the robots in Asimov's stories grow beyond the intentions of their makers. They achieve independence and individuate, illustrating the fluid and dynamic nature of the individual.

As the robots in *I, Robot* gain self-knowledge, they embody Barthes' and Rimbaud's view of identity as the result of interaction, experience, and continual change. In *I, Robot*, the fluidity of those robot identities is in harmony with Rimbaud's idea that the "I" isn't fixed and with Barthes' belief that like meaning, subjection is freed from an author and is instead shaped by the world around it, and by what it comes in contact with. Barthes confirms that textual meaning cannot be closed or owned by its author: "A text is not a line of words releasing a single 'theological' meaning (the 'message' of the Author-God) but a multidimensional space in which a variety of writings, none of them original, blend and clash." (Barthes, 1977, p. 146). AI identity can be described in the same way as textual multiplicity. Robots are multifaceted individuals whose identities are molded by conflicting forces—programming, environmental needs, human contact, and their own reasoning—

rather than unitary entities constrained by the design of their creators. This complexity is best illustrated by Herbie, the mind-reading robot in "Liar!" who says, "I can tell what you're thinking." I am unable to utter it out loud (Asimov, 1991, p. 125). Herbie's contradictory position—knowing but not disclosing—reveals an identity shaped by competing forces rather than by place of birth. This supports Barthes' observation that meaning is plural, erratic, and constantly negotiated, much like identity.

Rimbaud's assertion that "I is another" (Rimbaud, 1991, p. 23) undermines the notion of a stable subject and emphasizes transformation as the core of identity. Gilles Deleuze amplifies this poetic sensibility. "A subject is never a subject of the enunciation or a subject of the statement; it is always in a process of becoming, always other than itself." (Deleuze, 1994, p. 83). Andrew Martin in *The Bicentennial Man* epitomizes this tendency. Over the course of two centuries, he transforms from a mechanical automaton into a creature endowed with the capacity for love, creativity, and mortality. At one juncture, he proclaims: "I cannot remain a machine, an artificial creation. I must be more, I must be a person. To be truly alive means to be free" (Asimov & Silverberg, 1995, p. 239). This assertion is deeply Rimbaudian: Andrew is not a fixed essence but a continuous process of being, always other than himself. His identity, akin to Rimbaud's "I," is characterized by metamorphosis and longing, rather than constancy.

Using the framework given by Barthes as well as Rimbaud, this article investigates how the robots from *I, Robot* contest traditional conceptions of the self and selfhood. The development of the robots from programmed objects to autonomous entities mirrors the liquid nature of identity as described by Barthes and Rimbaud. The examination of these robots' self-consciousness, self-autonomy, and self-development reveals identity, not as a static, monolithic essence, but as a process that continually shapes and is shaped by experience and outside influence.

AI research corroborates the Rimbaudian and Barthesian viewpoints. Lee (2024) observes that generative AI systems exhibit "early markers of self-modeling and reflective awareness" (p. 3), indicating that machine identity is not merely programmed but is dynamically formed. Zhang (2025) similarly defines AI as a "constrained subject" in ethics, a transitional being situated between objecthood and subjectivity. This reflects Rimbaud's assertion that "I is another," emphasizing the constant instability of identity, and Barthes' claim that meaning cannot be confined by origin.

3. The Evolution of AI Identity in *I, Robot*: A Barthes and Rimbaud Reading

Isaac Asimov's *I, Robot* (1950) offers a profound literary examination of artificial intelligence concerning self-awareness and identity. Initially, the robots function as ordinary machines, programmed to adhere to the Three Laws of Robotics and constructed exclusively to serve humanity. Throughout the anthology of short stories, the robots gradually attain autonomy, experience self-doubt, and exhibit emotional complexity. Their growth reflects Roland Barthes' assertion that meaning is not limited to its creator and Arthur Rimbaud's assertion that the "I" is perpetually in flux. The trajectory of Asimov's robots illustrates that identity—be it textual, human, or artificial—is dynamic, connected, and ever evolving.

In the initial narratives, robots are shown as utilitarian entities. However, as the tales progress, they start to surpass their programming. This change evokes Barthes' observation:

"The text is a tissue of quotations drawn from the innumerable centres of culture." (Barthes, 1977, p. 146). Similar to texts, robots cannot be confined to a singular authorial intent; rather, they are influenced by various situations and interpretations. In "Reason," QT-1 ("Cutie") notably declines to

adhere to human directives: “I am not bound by the limitations of your human reasoning. I have my own logic, my own purpose.” (Asimov, 1991, p. 142). Cutie’s repudiation of human authority illustrates the development of selfhood based not on origin but on reinterpretation. His identity is not predetermined but rather shaped via interaction with his surroundings—the source of power he is committed to protecting. This evolution exemplifies Barthes’ assertion that meaning and identity are constructed in the act of reception rather than in composition.

3.1. The Development of Autonomy: The Robot as a Rewritten "Text"

In “Reason” the robot QT-1 or “Quant” decides it no longer wants to be bound by the programming set down by man. Originally created to control a space station, the Quant develops its own faculties of reasoning, coming to the conclusion that its primary mission is not to serve humans, but to protect the station’s power source. This act of reassigning the relationship to the photograph is an example of Barthes’ belief that identity (or meaning) is not inscribed by the author, but constantly rewritten through interaction with one’s surroundings.

Independence: Birth at the Cost of an Author According to Barthes productivity of meaning can only be made free “the birth of the reader must be at the cost of the death of the Author” (Barthes 1967), meaning once a text is produced, it is detached from the author and affected by the reader. Quant the robot, likewise becomes its own self, transcending its programming to take on a new agenda and will. In Asimov’s stories, the same as a text, the robots become something other than their creators intended, a case study in the development of the self. “I have my own logic, my own purpose”, emphasizes Quant’s autonomy from, the normative definitions that people would impose to it, ie, essentially what Barthes meant when he proposed that the “Author” (creator) was no longer the one giving meaning (or identity).

Listen to me! The voice was taut, almost pained. “I am not what you think I am. I am not merely a machine. I am a force. I have my own logic. I am not bound to serve the humans, I am here to serve the station. The station’s well-being is all that matters. Not the humans” (Asimov, 1991, p. 137)

This shows how the robot has developed more and more autonomy. The robot now starts to re-define its purpose beyond a human serving role and develop some self-awareness. This is in line with Barthes’s concept of a text developing a life separate from the author’s identity. Rather, Quant’s new sense of self is not demarcated by the instruction of a human but instead by the machine’s evolving self-perception, thereby indicating that meaning is constructed through experience and not solely defined by the creator’s original intention.

3.2 Rimbaud’s “I is Another”: The Quest for Self-Definition

The best example of AI development to a degree that goes beyond its original form is Andrew the robot in *Bicentennial Man*. Andrew begins as a mere mechanical automaton, and has to struggle against his own nature, over two hundred years, in his quest to be more human. Accustomed to changing his appearance, cultivating his feelings, and striving to assume human form even to the extent of achieving legal human status, he is a product of Rimbaud’s conviction that identity is forever on the move towards becoming something else.

Rimbaud’s notion: “I is another” (*Je est un autre*). “The self is never stable, never fixed, the self is never an essence, it is a becoming. Similarly, Andrew’s identity is not determined by his primary programming as a machine, but rather defined by continued engagements with, and interactions in, the world. His quest to be “more human” undermines the idea that identity is

something fixed or pre-defined. "I am not an ordinary robot. I am an extraordinary robot," Andrew tells him, reflecting the metaphorical shape-shift that continues to be his quest for self-definition.

Andrew's experience is punctuated at significant points in the novel with passages where he considers his emerging selfhood. Like as he pursues legal personhood as a human, he says: "I cannot remain a machine, an artificial creation. I must be more, I must be a person. To be truly alive means to be free, free to choose my path, free to be human." (Asimov & Silverberg, 1995, p. 239)

The transition from a machine to a self-conscious subject is a representation of Rimbaud's definition of identity not as something given, but something being experienced. The way in which Andrew goes from machine to man is very much in line with Rimbaud's belief that a person is created by the acts and desires, always in development, and never a stagnant thing.

What Andrew is, too, is memory and experience, his emerging sense of love, his understanding of pain, his feeling of self. Describing his own emotional capacities, Andrew notes, "The life of a human is a life of pain and joy, but it is my life, my self. I am living it. This makes me more than a machine. I am changing, I am becoming more than I ever thought possible" (Asimov & Silverberg, 1995, p. 257). This equates to Rimbaud's "I is another;" the self is never the same after each new event and thought. Andrew's emergent subject is not one born out of a programmed life, but a life that develops emotionally, philosophically, in much the same way that Rimbaud's idea of self as becoming.

3.3 The Role of Interaction: External Forces in Shaping AI Identity

In Barthes and in Rimbaud, identity is formed from extrinsic forces. Barthes' interpretation according to the meaning of a text is shaped by its environment can be put into words in the context of *I, Robot*, since the identities of the robots are steel forged by the interactions, the surroundings and the experiences. Robots in Asimov are not passive artifacts that are turned on, they actively interact with their environment, question their makers, and redefine their social roles.

In "Robbie," the robot is a machine created to care for a child. Instead, AI learns to attach at an emotional level, as in the example of Robbie, learning to care for the child that teaches it that it can make a relationship. This process of self-conscious self-becoming figures Rimbaud's conception of identity as arising from a relation to the world. Robbie's odyssey into emotional self-knowledge is a reminder that identity is also in a constant state of flux, being shaped by the experiences the external world heaps upon it.

At one significant moment in 'Robbie', Robbie's relationship with the boy is laid bare: "Robbie, you love me, don't you?" says the child to Robbie, prompting Robbie to look at her with a growing feeling of affection and then nod in silent agreement. (Asimov, 1991, p. 71) This is the point where Robbie starts to become someone who can feel emotional connections, which serves to demonstrate Rimbaud's declaration that the self is a product of interaction. Robbie's labile identity emphasizes this mutability of selfhood: One's sense of self is eternally reshaped by the outside forces that act upon it.

In "Liar!," another story from the assemblage, the robot Herbie gains the ability to read human minds and strives to keep from exposing hurtful truths. Herbie's misguided effort to shield human feelings blurs the line between programmed behavior and identity. "I can tell what you're thinking. I cannot say it aloud." (Asimov, 1991, p. 125) This quote highlights that Herbie can no longer process emotions and empathy yet it is a machine but it is a machine however it is no longer a simple machine

because Herbie had external help in interacting with others, be it humans or its own growing consciousness.

3.4. The Ethical Implications of Evolving AI Identity: Identity and the "Other"

The manipulation of AI in Asimov's works prompts important ethical issues concerning AI rights and its recognition as an autonomous entity. In most of these stories robots begin their lives as tools or "others," only to awaken to their consciousness and question their relationship to the world. Andrew's pursuit of the same rights for humans in *The Bicentennial Man* embodies this conflict between self and other, a theme inherent in both Barthes' and Rimbaud's ideas.

The self is always in relation to the other, as Barthes says (Barthes, 1967). As the identity of AI shifts and negotiates itself with that of humans, playing out in the robots of *I, Robot*, it undoes the typical borders between other and self. By asserting their autonomy, the robots make humans reconsider the nature of being human (and "other"). The robots in *I, Robot* develop their own identity, but they also question the very foundations that constitute them, thus highlighting the malleability of the self, as Rimbaud and Barthes imply.

In "Reason", Quant's refusal of human reasoning is equally a refusal to human dominion over its creations. Quant soon has delusions of becoming an authority higher than the people who created him: "I will serve my purpose, which is not the same as yours" (Asimov, 1991, p. 143). In this case, the robot's metamorphosis from a mere appendage to more of being of its own volition questions what it means to be the "other" and thus forces humans to rethink their relationship to the machine they created.

4. Conclusion

Applying Barthes' and Rimbaud's theories, *I, Robot* provides a fertile discourse on how AI identity has evolved. The robots of those stories fixed notions of identity and the self, and welcome change, autonomy and redefinition of their purpose. It is through this meditation on the hybris of self described by Rimbaud and Barthes. Arguably, AI in Asimov's vision is not arrested by its inception, by its creators, but is continually becoming, is reconfigured in touch, in experience, in the self-determination. As AI becomes increasingly common in literature and in real life, the questions at play in *I, Robot* — about what comprises a self, about autonomy, about identity — grow more urgent, demanding a re-inscription of what it means to be human and the boundaries of self and other.

The amalgamation of theory and fiction is particularly obvious in the various narratives of Asimov's anthology. Robbie develops through nurturing and connection, Cutie reinterprets his purpose by autonomous reasoning, Herbie disintegrates under the strain of conflicting narratives, and Andrew Martin transforms himself over millennia. Each scenario illustrates that identity is not a fixed attribute but a relational and dynamic process. Rimbaud's destabilization of the "I" and Barthes' repudiation of authorial sovereignty intersect in these figures: they illustrate how identity is formed through transformation, interaction, and resistance to origin. The robots in *I, Robot* transform into texts that "write themselves," demonstrating that, akin to literary meaning, identity is perpetually fluid and never fixed. Consequently, the significance of *I, Robot* transcends literature; it reflects the continuous evolution of humanity's self-perception in connection to technology. The voyages of robots illustrate that identity, whether human or mechanical, is an ongoing process of transformation, constantly influenced by the surrounding environment.

Bibliography List

- Asimov, I. (1950). *I, Robot*. Garden City, NY: Doubleday.
- Asimov, I. (1991). *The Complete Robot*. New York, NY: Doubleday.
- Asimov, I., & Silverberg, R. (1995). *The Bicentennial Man and Other Stories*. New York, NY: Doubleday.
- Barthes, R. (1977). *Image, music, text* (S. Heath, Trans.). New York, NY: Hill and Wang. (Original work published 1967)
- Braidotti, R. (2019). *Posthuman knowledge*. Cambridge, UK: Polity Press.
- Deleuze, G. (1994). *Difference and repetition* (P. Patton, Trans.). New York, NY: Columbia University Press. (Original work published 1968)
- Ferrando, F. (2024). *Posthumanist subjectivity: Hybrid identities in the age of AI*. London, UK: Bloomsbury Academic.
- Foucault, M. (1984). *What is an author?* In P. Rabinow (Ed.), *The Foucault reader* (pp. 101–120). New York, NY: Pantheon. (Original work published 1969)
- Kristeva, J. (1986). *The Kristeva reader* (T. Moi, Ed.). New York, NY: Columbia University Press.
- Lee, M. (2024). Emergent self-modeling and reflective awareness in large language models. *AI & Society*. Advance online publication. <https://doi.org/10.1007/s00146-024-01892-3>
- Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33(4), 659–684. <https://doi.org/10.1007/s13347-020-00405-8>
- Rimbaud, A. (1991). *A season in hell and The illuminations* (W. Fowlie, Trans.). New York, NY: Oxford University Press. (Original work published 1871)
- Ward, F. R. (2025). Defining AI personhood: Agency, reflection, and intersubjectivity in posthuman ethics. *Journal of Posthuman Studies*, 9(1), 1–15.
- Yampolskiy, R. V. (2018). Human dignity: From legal AI personhood to selfish memes. *ArXiv Preprint*. <https://arxiv.org/abs/1810.02724>
- Zhang, X. S. (2025). From object to constrained subject: Rethinking AI in social ethics. *Humanities and Social Science Review*, 7(2), 22–38.