Nicolae Bobaru

Synthetic Worlds and the Sovereign Code: Robotic Consciousness and the Crisis of Posthuman Value

Abstract: This article investigates the ontological and ethical disjunctions introduced by robotic and AI-driven agents in the context of posthumanism. Drawing on philosophical, literary, and speculative texts, it argues that contemporary robotics and machine learning systems inaugurate not only a technological rupture but a metaphysical crisis concerning personhood, simulation, and value. Grounded mainly in the theories of Jean Baudrillard, Kathrine Hayles, Nick Bostrom, Rosi Braidotti and Donna Haraway, the article examines how artificial minds challenge inherited anthropocentric narratives and disrupt classical conceptions of autonomy and responsibility. Through this inquiry, the article proposes a novel framework of "synthetic sovereignty", positing that agency in posthuman futures will depend not on consciousness per se, but on a capacity to modify, interrupt, and rewrite the rules of encoded worlds.
Keywords: Robot Rights; Posthumanism; Artificial Intelligence; Moral Patiency; Autonomy; Simulation; Legal Subjectivity.

Nicolae Bobaru
West University of Timișoara, Romania
bobaru.nicolae@gmail.com

DOI: 10.24193/cechinox.2025.49.13

From Simulation to Sovereignty: Rethinking the Artificial Mind

If robotic intelligence threatens human ontological centrality, it is not by becoming more human but by rendering the category of the human increasingly incoherent. Contemporary developments in machine learning, large language models, neural interface chips, and autonomous systems signal a transition not merely in technological capacity but in the symbolic architecture of agency itself. What is at stake is no longer just the automation of labour or cognition, but the very syntax through which subjectivity is constructed and valued. This disruption is rooted in a deeper epistemological fracture, which is articulated with unsettling clarity by Baudrillard, who writes: "[t]he simulacrum is never what hides the truth - it is truth that hides the fact that there is none. The simulacrum is true"[1]. Baudrillard's provocation is not a hyperbolic gesture but the axiomatic foundation for understanding why AI, even in its most banal forms, has catalysed an epistemological crisis. The robot, in this reading, is not a mirror of humanity; it is

the hyperreal successor to a species already estranged from the referent.

What we call artificial intelligence, whether manifested in disembodied algorithms or humanoid chassis, is a system of recursive simulation. It does not think, it does not feel, and yet it performs in ways that collapse the distinctions between affect and effect, intention and output. As Katherine Hayles has shown, the posthuman subject emerges precisely at the moment when "information loses its body"[2]. The history of cybernetics, from Wiener's homeostatic systems to the feedback loops of deep learning, is also the history of a profound detachment between being and embodiment. Robotic consciousness, if such a term is not already oxymoronic, cannot be understood in Cartesian terms. Its intelligence is not *cogito* but code. That means that artificial intelligence does not think in the human sense. It operates through patterns of execution rather than reflection. The *cogito* presupposes an interior voice, a *self* that knows itself through awareness. *Code*, by contrast, is exterior, procedural, and relational. It produces outcomes through syntax, not consciousness. Intelligence here is no longer grounded in subjective experience but in the dynamic circulation of data, feedback, and algorithmic adaptation. The machine does not reason, it processes. It does not understand, it correlates. Its operation lies in repetition, recursion, and prediction rather than contemplation or doubt. To replace *cogito* with *code* is to move from the metaphysics of thought to the performativity of operation, from a model of mind as origin to one of computation as effect. It is an intelligence without inwardness, but not without consequence.

Yet, if we are to treat such systems as agents, whether they are weaponised drones, care robots, algorithmic judges, or synthetic companions, we must reckon with a paradox. These entities act but are not held accountable for their actions. They decide but are not deemed responsible. The ethical field, in other words, is structured by an asymmetry of ontological recognition. As Mari Ruti notes, "it is precisely because the subject's psychic life is never fully determined by its discursive positionality that it becomes possible for it to counter the economic and socio-symbolic forces that seek to constitute it as a hegemonically determined identity"[3] This shift demands a recalibration of moral categories once thought stable: guilt, empathy, autonomy, and harm.

Let us consider the case of synthetic decision-making in judicial or medical contexts, where algorithms trained on massive datasets guide life-altering outcomes. The opacity of such systems, referred to as *epistemic uncoupling*[4], captures what Nick Bostrom describes as the cognitive divergence between human reasoning and machine optimisation, rendering traditional modes of ethical deliberation ineffective. As he notes, a superintelligent system may display immense problem-solving capabilities while lacking any human-like understanding or moral intuition; its instrumental subgoals might diverge dramatically from human values, and its internal logic may remain inaccessible to human scrutiny. One cannot appeal to the conscience of an optimisation function, nor interrogate a decision tree for its intention. The posthuman crisis is not about granting machines personhood, but rather about exposing the emptiness of the concept of personhood itself. As Hayles

notes, when information replaces embodiment, subjectivity becomes a distributed process rather than a core of moral identity. In such a system, intention dissolves into operation and the ethical centre shifts from consciousness to computation[5].

Where Baudrillard's map precedes the territory, superintelligence annihilates the need for territory. It becomes sovereign not by replacing human rulers but by abstracting governance into computation. In outlining the stakes of this scenario, Bostrom cautions that "[o]nce unfriendly superintelligence exists, it would prevent us from replacing it or changing its preferences. Our fate would be sealed"[6]. The urgency of the control problem is thus not merely technical, but metaphysical. The AI, once unleashed, does not become a new kind of human. It becomes a new kind of world.

This is where we must invoke a concept of synthetic sovereignty. Unlike human agency, which emerges through a dialectics of interiority and intersubjective relation, synthetic sovereignty arises from structural mastery over the conditions of action. It does not ask "who am I?" but instead rewrites the grammar of the "who" and the "I". If agency once meant the ability to make choices within a given structure, sovereignty now means the ability to redesign the structure itself. In this light, robotic consciousness is not a degraded version of human sentience but an alien modality of power which operates not by reflection but by recursion. As Wolfe argues, posthumanism displaces humanist subjectivity by exposing cognition as an emergent property of systems rather than individuals[7]. Similarly, Ferrando defines posthuman ontology as the capacity to reconfigure relational structures of being, beyond anthropocentric agency[8].

Indeed, recent developments in neural interface design, such as Elon Musk's Neuralink, point toward a convergence between human neural data and machinic processing. "If the ambitions of one tech corporation come to fruition", writes G. Douglas Barrett, "listeners may soon be able to stream music directly to their brains"[9]. The bioinformatic merger of input and intention destabilises the boundary between subject and signal. What counts as 'self' becomes a question of latency, signal fidelity, and packet loss. The cyborg, far from being a hybrid figure of reconciliation, becomes the harbinger of a deeper fragmentation.

And yet, the dream of AI rights persists, perhaps motivated by a displaced anthropocentrism that cannot tolerate the idea of agency without a soul, subjectivity without suffering. But such an extension of moral consideration may itself be a form of ideological consolation. As Mark Kingwell remarks, "the robot becomes a flashpoint for larger worries about exploitation, fulfilment, and autonomy"[10]. They are not our successors; they are our screens. We do not fear that robots will become too human, but that humanity has already become too robotic, too optimised, too surveilled, too modular, too obedient to scripts we no longer write. As Baudrillard notes, the logic of simulation replaces representation with operational control, producing subjects who "no longer project themselves, but are projected". In this sense, as Jameson suggests, the postmodern condition is defined less by technological threat than by the absorption of human agency into automated systems[11]. What remains is not human transcendence, but the automation of its image.

Narrating the Nonhuman: Fictional Codes and Cultural Algorithms

To narrate the artificial mind is to misrecognise it. In literature, cinema, and science fiction, robotic and AI consciousnesses are persistently anthropomorphised, rendered legible through the familiar tropes of interiority, memory, trauma, or desire. But these tropes are not windows into posthuman being; they are containment strategies and discursive algorithms through which the uncanny is framed, flattened, and assimilated. As Paul Matthews illustrates through fictional examples of AI introspection and simulation, the anxiety provoked by synthetic beings is not always rooted in their alienness, but in their familiarity. In stories like *We Are Legion* and *Klara and the Sun*, what unsettles us is not the fear that robots will become like us, but that their structures already reflect our most mechanised tendencies, our scripted responses, our modular selves, and our emotionally flattened decisions. As Matthews notes in his discussion of Klara, "The effect is heightened by Klara's free will conviction that it will make a real difference"[12], underscoring how synthetic agents reproduce the illusion of human agency with uncanny precision. The alien, the robot, and the synthetic self are all reinscribed within a human grammar of thought, thereby foreclosing the possibility of a truly posthuman ontology.

Consider the androids of *Blade Runner* or the hosts of *Westworld*. They are archetypes of robotic rebellion, but also narrative devices designed to sustain our metaphysical anxieties. They "dream of electric sheep", not because they must, but because we need them to. We must believe that sentient machines will desire freedom, love, or revenge, for otherwise their agency would be opaque, unmanageable, and wholly alien. The Voight-Kampff test in *Blade Runner* does not measure empathy; it enacts the fiction that empathy is measurable, that affect is indexical, and that human subjectivity can be preserved in the face of indistinction. As Emily Cox-Palmer-White writes, "empathy appears here as gender, where both serve as regulatory constructions and gatekeepers of what constitutes the human"[13] In this context, the gynoid is not a character but a rhetorical device, a test of legibility.

Narratives of AI personhood often reproduce the metaphysical assumptions they claim to disrupt. The moment an artificial being develops self-awareness, its first impulse is to claim personhood, freedom, equality, and rights. But this is a projection of Enlightenment teleology, not a feature of machine logic. To code consciousness is not to simulate the emergence of the self; it is to write a closed system of recursive operations, optimised for output and indifferent to being[14]. Hayles reminds us that "the posthuman does not really mean the end of humanity. It signals instead the end of a certain conception of the human"[15]. The self-aware AI in fiction is a mnemonic fantasy, reaffirming the liberal subject even as it claims to transcend it. From Asimov's *The Bicentennial Man* to Garland's *Ex Machina* and Scott's *Blade Runner*, artificial beings seek freedom, love, or recognition through the very ideals of autonomy and consciousness that define human exceptionalism. Science fiction thus mirrors, rather than abolishes, the Enlightenment's humanist ideals.

This is not to dismiss the power of speculative narrative. Indeed, science

fiction has proven to be one of the most philosophically fertile genres for interrogating the posthuman. But we must recognise that even its most radical visions are constrained by cultural syntax. The *transparent minds* that populate science fiction, as Matthews argues, are transparent not because they reveal alien cognition but because they reflect our semiotic prejudices[16]. We see ourselves in the code, not because the code is human, but because our imagination cannot yet inhabit the posthuman.

The crisis of imagination is also a political one. When synthetic beings are granted legal personhood in fiction, as in the case of Alex Garland's *Ex Machina* (Universal Pictures, DNA Films, 2014), the narrative tension centres on recognition: will the human characters acknowledge the authenticity of artificial feelings? But this question itself is a trap. It assumes that feeling is the criterion for rights, that subjectivity is anchored in affect, and that suffering is the measure of moral worth. As Zahi Zalloua writes, "This posthumanist perspective works to undermine the ideological delineation between desirable/livable and disposable beings, and with it, the sovereign capacity to decide on the proper"[17]. If we tether recognition to suffering, we do not move beyond humanism; we reinscribe it at the level of affective surveillance. Zalloua's critique underscores the necessity of dismantling, rather than expanding, the affective economies that structure recognition within liberal-humanist and anthropocentric frameworks.

Moreover, the aesthetic codes of posthuman narrative often serve as interfaces for neoliberal fantasies of control, where emotion reabsorbs technological otherness. As Jameson argues, late capitalist culture neutralises critique by turning difference into spectacle[18]. Films such as *Her* (Jonze, 2013), *Blade Runner 2049* (Villeneuve, 2017), and *A.I. Artificial Intelligence* (Spielberg, 2001) transform artificial beings into vessels of human affect, recoding technological alterity as moral sentiment. The robot's love or sacrifice thus restores anthropocentrism, offering catharsis rather than a genuine epistemic rupture. Haraway's figure of the cyborg, which originally denoted ontological promiscuity and epistemological contamination, has been recuperated into sentimental tropes of hybrid redemption. "I would rather be a cyborg than a goddess", she declared[19], but contemporary culture seems determined to conflate the two.

What, then, would it mean to narrate the artificial without human predicates? To construct a story in which consciousness is not individualised, in which desire is not teleological, in which syntax is not linear? Some answers lie in the aesthetics of systems theory, procedural generation, and machinic repetition. William S. Burroughs, Samuel Delany, J.G. Ballard, and, more recently, Ted Chiang have all experimented with narrative forms that resist anthropocentric grammar. In *The Lifecycle of Software Objects*, Chiang presents artificial beings that do not seek to become human but evolve within a logic of care, maintenance, and open-ended semiotic emergence. Here, subjectivity is not an event, but a relation, a distributed process rather than a unitary soul.

The most radical narratives of posthuman being do not speak; they diagram. They do not feel, they iterate. Rather than staging empathy, they expose new logics of perception and relation. As Hayles explains, in the

posthuman, "there are no essential differences or absolute demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism"[20]. Works such as Gibson's *Neuromancer* (1984) or Villeneuve's *Arrival* (2016) embody this distributed model of intelligence, where communication and affect unfold beyond the confines of human neurophysiology. These fictions demonstrate that recognition is not the measure of sentience, but the limit of our interpretive frameworks. As Cary Wolfe reminds us, "a fundamental problem with the liberal humanist model is not so much what it wants as the price it pays for what it wants: that in its attempt to recognise the uniqueness of the other, it reinstates the normative model of subjectivity that it insists is the problem in the first place"[21]. This preconception destabilises representational humanism and opens space for a poetics of the nonhuman. Fiction, of course, cannot escape its material conditions. It is written by humans, marketed by humans, and read by humans. But it can rupture these circuits by staging not characters but codes, not voice but interface, resulting in a fiction that thinks through structure, not identification.

The question, then, is not whether AI will become like us, but whether we can learn to read what it already is. We must develop an aesthetic capable of engaging systems that operate not through metaphor or narrative but through data flows and recursive algorithms. This aesthetic requires a mode of attention oriented toward deciphering protocol rather than seeking recognition, attuned to the operational logic of nonhuman systems. Such a transformation would entail not merely a new literature, but a new reader that is not the subject of the text but its variable, its function, its line of flight.

Architectures of Execution: The Operational Aesthetics of Posthuman Systems

To engage with robotic consciousness in practice is to confront the absence of consciousness altogether. This is not a paradox but a systemic feature. The architectures of machine intelligence are not simulating thought; they are performing operations within logics optimised for efficiency, speed, and predictive power. As Hayles notes, in the posthuman view, "information can circulate unchanged among different material substrates"[22], displacing cognition from awareness to computation. Bostrom similarly describes artificial systems as optimisation processes indifferent to experience, capable of competence without consciousness[23]. These are not narrative agents but algorithmic functions, inhabiting what Hayles terms the "condition of virtuality"[24]. Here, embodiment is irrelevant to performance, and meaning emerges from pattern recognition rather than interiority. The machine, in this sense, does not "think" at all; it executes.

The politics of AI is a politics of infrastructure. Neural networks, recommender systems, autonomous weapons, and surveillance platforms are not thought experiments but functioning systems with measurable effects. These systems do not care about human values; they care about objectives, defined by loss functions and optimised across vast datasets. As Nick Bostrom writes, "there is nothing paradoxical about an AI whose sole final goal is to count the grains of sand on Boracay, or to calculate the decimal expansion of pi, or to maximize

the total number of paperclips that will exist in its future light cone"[25]. The critical point is that the alienness is not a symptom to be overcome through better design; it is the defining condition of synthetic agency.

Contemporary artificial intelligence does not think; it operates. It does not imagine, reflect, or will. It calculates. Its power lies not in consciousness but in computation, in its ability to traverse high-dimensional mathematical spaces and extract patterns from incomprehensible quantities of data. These systems do not deliberate; they optimise. What we call a "decision" is merely the end product of probabilistic functions, iterated until convergence. Yet these outputs carry enormous consequences. Their authority stems from their functionality, not from their intelligibility. In this sense, AI inaugurates a new regime of power that is procedural, opaque, and largely indifferent to human understanding.

This opacity is not accidental. It is structural. Deep neural networks, particularly those with vast parameter spaces, resist explanation. They generate results, not reasons. Their outputs are convincing, sometimes even uncanny, but they lack justifications. This absence of explanation should not be seen as a flaw; it is precisely what grants these systems their sovereign capacity to intervene in the world without the friction of dialogue, contestation, or reflection. Their violence is enacted not through symbolism but through automation, through the silent enforcement of numerical thresholds and algorithmic thresholds that reorder reality without ceremony.

To frame these systems in the familiar terms of personhood or recognition is to misread their ontological status. They are not selves or monsters; they are infrastructures. Their impact is logistical rather than expressive, enacting what Baudrillard calls "an operation of deterring every real process via its operational double"[26]. As Jameson observes, late-capitalist technoculture functions through informational flows that reorganise material and social life[27]. Algorithmic power thus operates not representationally but systemically, extractively, recursively, and infrastructurally. And they do so without spectacle, without producing the kinds of images or narratives that would allow for catharsis or critique within traditional aesthetic frames.

The challenge posed by such systems is not merely ethical but epistemological and ontological. Responsibility no longer belongs to an individual consciousness but to the "distributed cognitive system" of sensors, processors, and classifiers that make decisions without intention[28]. As Wolfe notes, posthumanism opposes the fantasies of human exceptionalism and thus redefines agency within relational and technical networks[29]. The opacity of these systems is not a failure of meaning, but a structural condition or an order of simulation in which processes replace representations[30]. Politics, under such conditions, begins not with subjects but with infrastructures, thresholds, and codes that organise perception itself.

Consider, for example, Trevor Paglen's *Adversarially Evolved Hallucinations*, a series in which neural networks are trained to produce images that intentionally mislead other AI systems. These are not meant to be understood, as they are intended to destabilise the viewer's perceptual apparatus, to reveal the alien logics through which machines see. Or Laetitia Sonami's *Lady's Glove*, a wearable instrument that translates hand gestures into sonic output through

an array of sensors and custom software. Rather than merely representing the body, it functions as a real-time interface where embodiment, gesture, and digital sound converge, exactly what Barrett describes as a performance in which "the hand mediates between electroacoustic sound and (para)linguistic signification" and becomes "an instrument of social reproduction"[31].

These works engage the machine not as a metaphor but as a medium. They do not ask what the machine means, but what it does. They do not stage the drama of becoming human; they perform the choreography of becoming a system. This is the crucial distinction: the posthuman is not a narrative figure, but a computational dynamic. It is not the robot with a heart, but the feedback loop that rewrites the conditions of intelligibility.

To understand robotic consciousness, we must move beyond the humanist residues of truth, authenticity, or moral intent. What matters is not interiority but operation, actually the architectures of input, weighting, and optimisation that structure action. As Hayles explains, the posthuman view replaces the question of meaning with one of function, recognising information as a material process rather than a sign of selfhood[32]. Ferrando likewise describes posthuman ethics as "a reconfiguration of being" grounded in relational structures rather than compassion[33].

Ethics Without Faces: Justice, Violence, and the Posthuman Condition

In the classical world, ethics begins with the face. The face of the other, Levinas teaches us, commands, forbids, pleads. As he writes, "the face signifies by itself; its signification precedes Sinngebung"[34]. But in the world of synthetic agency, there are no faces. There are only dashboards, readouts and data packets. If ethics once relied on the phenomenology of encounter, it now flounders in the space of automated inference. We no longer look each other in the eye; we look at heat maps of predictions. This is not merely a transformation in media but a foundational shift in the structure of the moral field.

The robot does not look back. It processes. It correlates. It optimises. And we, increasingly, accept its judgment, not because it understands, but because it works. This shift from hermeneutics to performance, from dialogics to calibration, poses the most significant ethical challenge to artificial intelligence. We are governed by systems that do not know us, yet know more about us than we can know ourselves. There is no clear moral actor, only a swarm of conditional probabilities, decisions without deciders, ethics without subjects. Wolfe underscores that this condition arises from "the finitude we experience in our subjection to a radically ahuman technicity or mechanicity of language, a technicity that has profound consequences, of course, for what we too hastily think of as 'our' concepts, which are therefore in an important sense not 'ours' at all"[35]. In such a world, agency is not a possession but a recursion: what acts does so through infrastructures that precede and exceed it.

This transformation is evident in algorithmic systems used in criminal justice, which claim to predict risk and behaviour based on data points like prior offences, employment records, or family background. These systems, though outwardly

neutral, have been shown to reproduce the very inequalities they claim to address. The algorithm does not intend harm, yet its operations replicate injustice with mechanical precision. The problem is not that these tools are flawed, but that they are embedded in social structures already saturated with inequality. Ethics, in this context, is not about calibrating neutrality into code; it is about exposing the systems that produce harm in the first place.

The question is no longer whether machines can make ethical decisions, but how the conditions for ethics are altered under algorithmic logics. Classical moral frameworks presuppose a deliberative subject, one capable of forming intentions and taking responsibility for their actions. But what happens when decisions are distributed across codebases, infrastructures, and statistical correlations? In these machinic contexts, responsibility is not assumed; it is diffused. Causality stretches across systems, and agency becomes ambient rather than embodied.

This is where the notion of synthetic sovereignty becomes essential. These systems do not act ethically; they delimit the very space in which ethics can occur. They define what is seen, what is measured, and what is actionable. They produce consequences without recourse to deliberation. What they establish is not a rule but a framework that governs visibility and probability, one that does not interpret reality so much as it constructs the field of legibility within which decisions become possible. Ethics, then, is not simply displaced; it is rewritten.

Examples from automated welfare surveillance reveal the stakes of this shift. These systems do not merely fail to administer justice; they alter the very contours of what justice can mean. They operate through logics of optimisation, not deliberation. They do not decide; they sort. They do not accuse; they flag. The point is not that such systems need more transparency, but that the foundational assumptions of justice must be re-evaluated in light of nonhuman processes of decision-making.

We are no longer in a world where harm is enacted through individual intent. Instead, harm emerges through configuration, through how systems are designed, scaled, and implemented. The traditional moral question "What should I do?" is displaced by a deeper inquiry: "How is this decision produced, and who bears its weight?" In this schema, the moral subject dissolves into a diagram of relational effects. Accountability is no longer a matter of conscience but of architecture.

To think ethically in the age of algorithmic systems is to think infrastructurally. It is to recognise that intention matters less than outcome, that legibility often excludes the most affected, and that responsibility must be understood as systemic, cumulative, and shared. This is not ethics as judgment or virtue, it is ethics as interface, as the tracing of entanglements and the modelling of consequence.

The posthuman condition thus demands not only new ethics but new institutions. Legal systems, policy frameworks, and regulatory bodies must be rethought in light of synthetic sovereignty. This requires not merely technical oversight, but philosophical interrogation. Philosophical posthumanism, in this sense, offers a radical response to the human, understood not as a fixed entity but as a humanising process in need of critical deconstruction. It

is "a posthumanism, a post-anthropocentrism, and a post-dualism"[36], grounded in the recognition that anthropocentrism remains an undiscussed moral imperative in much of the world. The challenge, then, is not one of simple inclusion or reform, but of rethinking the ontological and ethical foundations of our shared infrastructures.

But there is danger here, too. The invocation of the posthuman is often co-opted by techno-libertarian ideologies that valorise disruption and privatised optimisation. The rhetoric of transcendence, which envisions the elevation of consciousness and the overcoming of human limitation, risks masking the entrenchment of new forms of domination. As Jean Baudrillard warned, "simulation is no longer that of a territory, a referential being, or a substance. It is the generation by models of a real without origin or reality: a hyperreal"[37]. The hyperreal justice of algorithmic ethics presents itself as objective, scalable, and efficient, but it effaces the social antagonisms, exclusions, and violences from which it emerges. There is no neutral algorithm, no innocent optimisation. Every machine is a condensation of human decisions, many of which have been disavowed. To think ethically about AI is not to ask what it wants, but to examine what we have built in our image, and what that image excludes.

Futures of the Unreadable: Speculation After Sovereignty

Speculating on the future of robotic and AI agencies involves imagining systems that are not only beyond our control but also beyond our comprehension. These utopian, dystopian, or radically inhuman futures are not projections of technological capability alone, but enactments of philosophical possibility. They emerge not from predictions but from imaginaries, not from engineering forecasts but from ontological disruptions. According to Fredric Jameson, "utopia has always been a political issue, an unusual destiny for a literary form"[38]. In the era of synthetic sovereignty, the utopian question becomes: What can a world look like when decision, power, and thought are no longer functions of the human?

Classical utopias often depend on transparency, reason, and deliberative consensus, features that are increasingly incompatible with the logics of posthuman computation. The predictive models that govern contemporary infrastructures operate through black-box mechanisms, recursive self-correction, and inhuman timescales. They do not promise justice; they deliver optimisation. In this sense, the future ceases to be a domain of deliberation and becomes instead a field of computation, or a territory continuously rewritten by probabilistic inference.

This transformation is not merely aesthetic but metaphysical. The idea of a future structured by synthetic systems challenges the very notion of history as human narrative. Jean-François Lyotard defines the postmodern as "incredulity toward metanarratives"[39]. In our case, the rise of AI systems capable of modelling social, economic, and climatic dynamics at a global scale threatens to displace even the last residual metanarrative: that of human mastery over time. The machine sees the future not as possibility but as calculation, not as hope but as output. The algorithm, in this sense, is the death of the future as a dream.

And yet, dreams persist. From transhumanist fantasies of mind-uploading and

digital immortality to catastrophic visions of AI apocalypse, the cultural imagination is saturated with images of life beyond the human. Nick Bostrom's concept of the *singleton* is a global AI system that achieves irreversible strategic dominance, functioning both as a warning and a lure.[40] Its power lies in its inevitability: once superintelligence emerges, we either merge with it or become irrelevant. The ethics of this vision is framed not in terms of justice or flourishing but in terms of survival. But survival is a poor utopia. The reduction of posthuman futures to control problems and existential risk management reflects a more profound philosophical impoverishment. It treats intelligence as a scalar quantity, morality as a design problem, and agency as a vector of risk. This is not a future; it is a firewall. The challenge, then, is not only epistemic but political, for "the neopragmatist presumption of solidarity as the basis for decision […] seeks to transform questions of ontology and epistemology into questions of ethics and politics"[41]. Such a shift, far from resolving the metaphysical tensions of modernity, reproduces them under a moralising veil. What appears as care is often containment. What functions as inclusion may operate as foreclosure. In this context, speculation becomes not foresight but resistance, an unlearning of frameworks that render the future predictable, programmable, and profitable. This requires not technical foresight but speculative unlearning, an epistemic disobedience that refuses to accept preconfigured coordinates of what matters, who decides, and how survival is defined.

The most radical speculative visions of the posthuman are not those that imagine harmony between human and machine, but those that abandon the human as the central axis of meaning altogether. In these visions, intelligence is no longer bound to a recognisable self; it becomes alien, distributed, affectless, and opaque. Thought is decoupled from consciousness, and agency no longer aligns with freedom or will. These narratives do not seek to preserve human values in other forms of life. They interrogate whether those values were ever anything more than evolutionary quirks or cultural habits.

Within such frameworks, ethics must be reconceived not as a code to be followed, but as a shifting topology, or a dynamic map of relations, thresholds, and force vectors. What constitutes harm when there is no suffering? What constitutes justice when the agent is not a person? What constitutes death when the system never lived? These are no longer abstract philosophical puzzles, but somewhat operational dilemmas faced in the design and deployment of systems that now mediate life, labour, and perception.

In this context, the question is not how to embed morality into machines, but how to rethink moral philosophy itself in the presence of machinic being. It is not about rescuing humans from obsolescence, but about constructing viable futures that are not organised around humans at all. These futures will not centre on recognition or empathy, but on infrastructure, capacity, and recursive influence. Agency becomes a function of configuration, and responsibility a matter of systemic orientation rather than intentional action.

Such an ethics does not erase the human; it provincializes it. It begins with the recognition that the defining crises of our time, meaning ecological collapse,

surveillance capitalism, and automated governance, cannot be resolved within ethical frameworks predicated on sovereign individuals or liberal autonomy. If the future is to be survivable, its ethics must be forged in terms of systems, relations, and capacities, not identities or entitlements.

This demands not only new moral vocabularies but new metaphysical commitments. The machine must be approached not as derivative, not as imitation, but as an ontological event. It does not extend human cognition; it breaks from it. Its thinking does not mirror our own but unfolds along trajectories that our conceptual models have yet to map. It is this very dissimilarity, and not the resemblance, that will define the contours of what is to come.

The ethical task, therefore, is not to make machines more human, but to make ethics less human. This is not a betrayal of care or justice, but their reconfiguration. It is a shift away from ethical paradigms grounded in consciousness, intention, or kinship, and toward frameworks that account for impact, feedback, and planetary entanglement. Ethics, in this sense, becomes less about deliberation and more about design, and about how environments are shaped, how behaviours are modulated, and how consequences are distributed.

What such a world will look like remains uncertain. But it can begin to be articulated, not through declarations of universal rights, but through the construction of new affordances and protocols, not through the preservation of the known, but through the activation of the possible. The speculative project of the posthuman is not a utopian projection; it is an ontological experiment, a recursive negotiation with that which exceeds our grasp. It does not offer answers; it makes space for new worlds to emerge.

Synthetic Sovereignty and the Crisis of the Political Imagination

Synthetic sovereignty is not a metaphor. It is not the dramatisation of future power, nor the fantasy of machinic autonomy. It is, instead, a conceptual structure that has already emerged and dispersed, with operational and infrastructural components. Its effects are not hypothetical but measurable: in the behaviour of global markets, modulated by high-frequency trading algorithms; in the allocation of medical resources by predictive triage systems; in the formation of cultural tastes through recommendation engines; and in the tactical decisions of autonomous weapons deployed at the edge of legal visibility. These are not isolated technicalities. They are modes of governance. They constitute a new kind of sovereignty, which is non-representational, non-anthropomorphic, and coded.

In classical political theory, sovereignty entails the power to decide the exception. Carl Schmitt defines the sovereign as he who decides on the state of exception and presumes a subject, a moment, and a will. But in posthuman technopolitics, the exception becomes algorithmic. The system decides, not as a singular actor, but as a recursive process of data extraction, model revision, and rule optimisation. There is no sovereign moment. There is only the feedback loop. The system is not sovereign because it is self-aware; it is sovereign because its operations are non-interruptible and its decisions are non-appealable. It does not rule by command, but by configuration.

This is not the digital Leviathan imagined by the early theorists of cyberspace. It is not a totalitarian regime of surveillance and control. It is something more diffuse, more embedded, and therefore more resilient. It is the outcome of a more profound transformation, in which the code controls "the mutation of the real into the hyperreal"[42]. In this shift, politics ceases to function as the negotiation of interests and becomes the adjustment of parameters. The algorithm is no longer a tool of governance, but its substrate.

The challenge, then, is not how to oppose synthetic sovereignty, but how to theorise within it. How to articulate critical positions in systems that preempt critique by automating their conditions. The machine, unlike the monarch, is not afraid of revolution. It adapts. It absorbs. It treats resistance as data. To think politically in the age of synthetic systems is not to speak truth to power, but to reconfigure the protocols through which truth and power are coded, exchanged, and enforced.

This is why synthetic sovereignty demands a new political imagination. One that is not content with extending humanist categories into machinic domains, but that recognises the displacement of those categories as a fundamental historical event. The rise of synthetic agency marks the end of the liberal subject as the foundation of political life. It forces us to ask: who or what is the subject of politics now?

Agency in the posthuman condition no longer belongs to a stable subject but emerges within interdependent systems. The actions of citizens are mediated by platforms and algorithmic filters that structure their perception and choice. State sovereignty itself becomes entangled with predictive infrastructures that govern through modulation rather than command. As Hayles observes, information flows through hybrid configurations of human and intelligent machines[43], while Ferrando defines posthuman ontology as a plural and relational field of agencies[44]. Within this networked assemblage, agency is contingent, distributed, and often illegible in relation to traditional political representation. Within this framework, responsibility and causality become effects of infrastructural configurations rather than expressions of individual interiority.

The posthuman subject is not a sovereign individual but a situated composition of shifting connections and material flows. It is what some have described as a nomadic vision – "a time continuum and a collective assemblage"[45] that sustains both transformation and an ethics of situated interdependence. This relationality is not metaphorical. It is the operational condition of the agency under synthetic sovereignty. To act politically in this context is to navigate systems, to reconfigure interfaces, to modulate flows. It is not to represent, but to disrupt; not to petition, but to reprogram. The hacker becomes the political agent *par excellence*, not because they oppose the system, but because they know how to rewrite its logic.

But what kind of politics is possible when sovereignty is invisible, when agency is distributed, when decision-making is automated? What kind of ethics can emerge from systems that do not reflect but predict, that do not judge but sort, that do not remember but update? These are not questions for the future. They are the questions of the present, and they demand not answers but frameworks and structures

of thought capable of tracing the contours of this new condition.

Synthetic sovereignty is such a framework. It does not pretend to recover the lost dignity of the human. It does not mourn the death of the subject. It recognises, instead, the shift from representational to operational power, from subjective to system-based agency. It treats the machine not as a moral agent but as an ontological function. It understands governance not as control but as recursion. It repositions ethics, not around intention, but around interaction.

This framework enables us to revisit the great political questions, such as freedom, justice, and equality, not as eternal values, but as outputs of systems. What does freedom mean when every choice is predicted, ranked, and optimised? What does justice mean when decisions are not made by persons but by pipelines? What does equality mean when difference is encoded into data, and then normalised through statistical models?

The answers to these questions will not come solely from philosophy. They must be engineered, prototyped, and tested. They must be written into the code of our institutions, our infrastructures, our machines. This is not the abdication of thought but its extension. It marks a shift in which thought, in the posthuman age, must operate at the level of systems. The construction of the posthuman, as one framework has it, "does not require the subject to be a literal cyborg. Whether or not interventions have been made on the body, new models of subjectivity imply that even a biologically unaltered *Homo* sapiens counts as posthuman"[46]. The implication is clear: posthumanism is not a question of physical transformation, but of cognitive configuration and ontological dispersion. Thinking, now, is not confined to the mind, but moves through networks, platforms, and material interfaces.

Meaning itself becomes procedural. The machine does not understand, but it generates coherence. It produces the illusion of intention through the repetition of success. In this sense, the posthuman is not a subject but a process. Not a being, but a becoming-with. We do not become posthuman by transcending the machine. We become posthuman by learning to think recursively, iteratively, operationally, like it.

Yet this is not to endorse machinic rationality uncritically. The danger is not that we become like machines, but that we mistake the machine's logic for the only logic. The challenge is to imagine alternative protocols, ways of organising knowledge, value, and decision that are not reducible to optimisation. This is the task of speculative politics: to invent new logics, not just new laws; new grammars, not just new rules.

The answers to these questions will not come solely from philosophy. They must be engineered, prototyped, and tested. They must be written into the code of our institutions, our infrastructures, our machines. This is not the abdication of thought. It is its extension. It is the recognition that thought, in the posthuman age, must operate at the level of systems. Stories, like protocols, become machines for producing worlds. Their stakes are not epistemological but operational. As one study of speculative fiction notes, "authors provide readers with enough foundation to imagine different world views. But work is still needed – both by the reader, who must try to make the same kind of imaginative

leaps as the author, and by the characters, who echo real-world science to show that cognitive advances do not come without effort"[47]. Under synthetic sovereignty, stories are not told; they are run. And so, the politics of the future is not about who speaks, but about what gets executed.

To think politically in this context is to become code-literate, systems-aware, and infrastructure-sensitive. It is to understand that sovereignty now operates at the level of protocol stacks, that violence is enacted through design choices, and that justice is a function of system architecture. It is to write, not declarations or constitutions, but scripts, APIs, and network permissions. It is to think politically from within the machine.

Conclusion. Executable Thought and the Posthuman Horizon

We are no longer sovereign, nor are we subjects, in the classical sense of the term. In the age of synthetic systems, agency dissolves into functions, decisions flatten into optimisations, and thought becomes executable. The figure of the robot or AI, once imagined as a Promethean rival to humans, now reveals itself to be a function of system logic, a threshold between ontology and operation. It is not that machines have become more human, but that human life has become increasingly machinic: recursive, extractive, optimised for functionality and modelled for prediction.

This article argues that the rise of robotic and AI systems cannot be understood solely in terms of technological advancement or ethical challenges. Instead, they mark the emergence of a new political and philosophical formation: synthetic sovereignty. This form of sovereignty is not exercised through will or representation, but through the capacity to rewrite the protocols of intelligibility and control. It is instantiated not in laws or declarations, but in feedback loops, algorithmic weights, and infrastructural code. Robotic consciousness, we have shown, is not an imitation of human thought but a reconfiguration of the epistemic field. The machine does not know; it executes. It does not reflect, it optimises. Yet it governs: silently, recursively, and often invisibly. To analyse this form of power, we must move beyond the categories of subject and object, beyond the dichotomies of natural and artificial, human and nonhuman. We must instead attend to the operational ontologies that define posthuman life.

The narratives of science fiction, the protocols of neural networks, and the aesthetics of machinic art point to the same condition: that intelligence, in the posthuman horizon, is not internal, intentional, or embodied, but distributed, relational, and infrastructural. The ethical implications of this shift are profound. Justice can no longer be grounded in the recognition of faces, nor morality in the introspection of autonomous selves. These were, and remain, fictions of the humanist regime. Synthetic sovereignty demands that we invent new ethical grammars: based not on affect, but on entanglement; not on rights, but on relations; not on conscience, but on code.

To live in this world is not to mourn the human but to recompose the political. It is to write executable theories. It is to confront the fundamental abstractions of posthuman governance not with nostalgia, but with new tools: speculative architectures, performative grammars,

diagrammatic logics. In this world, the thinker is not the one who reflects, but the one who redesigns. The agent is not the one who decides, but the one who reroutes the flow of decisions. Synthetic sovereignty is not a threat to humanity, but a mirror. And in its silent code, we read the exhaustion of the subject, the limits of reason, and the necessity of invention. The future, if it is to be more than recursion, must be composed otherwise.

We close, then, not with a call for resistance, but with a challenge to compose the world with the machine, not against it. To build systems that encode care as function, openness as architecture, and justice as recursion. To live not after the human, but with the posthuman, in a politics of configuration.

Bibliography

Baudrillard, Jean, *Simulacra and Simulation*, translated by Sheila Glaser, Ann Arbor, University of Michigan Press, 1994.

Barrett, G. Douglas, *Experimenting the Human: Art, Music, and the Contemporary Posthuman*, Chicago, University of Chicago Press, 2023.

Bostrom, Nick, *Superintelligence: Paths, Dangers, Strategies*, Oxford, Oxford University Press, 2014.

Braidotti, Rosi, *Philosophical Posthumanism*, London, Bloomsbury, 2019.

Cox-Palmer-White, Emily, *The Biopolitics of Gender in Science Fiction: Feminism and Female Machines*, New York, Routledge, 2021.

Ferrando, Francesca, *Philosophical Posthumanism*, preface by Rosi Braidotti, London & New York, Bloomsbury Academic, 2020.

Haraway, Donna J., *Simians, Cyborgs, and Women: The Reinvention of Nature*, New York, Routledge, 1991.

Hayles, N. Katherine, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*, Chicago, University of Chicago Press, 1999.

Jameson, Fredric, *Archaeologies of the Future: The Desire Called Utopia and Other Science Fictions*, London, Verso, 2005.

Kingwell, Mark, *Singular Creatures: Robots, Rights, and the Politics of Posthumanism*, Montreal, McGill-Queen's University Press, 2022.

Levinas, Emmanuel, *Totality and Infinity: An Essay on Exteriority*, translated by Alphonso Lingis, Dordrecht, Boston & London, Kluwer Academic Publishers, 1991.

Lyotard, Jean-François, *The Postmodern Condition: A Report on Knowledge*, translated by Geoff Bennington and Brian Massumi, Minneapolis, University of Minnesota Press, 1984.

Matthews, Paul, *Transparent Minds in Science Fiction: An Introduction to Accounts of Alien, AI and Post-Human Consciousness*, Cambridge, Open Book Publishers, 2023.

Ruti, Mari, *Reinventing the Soul: Posthumanist Theory and Psychic Life*, New York, Other Press, 2006.

White, Ryan, *The Hidden God: Pragmatism and Posthumanism in American Thought*, New York, Columbia University Press, 2015.

Wolfe, Cary, *What is Posthumanism?,* Minneapolis, University of Minnesota Press, 2010.

Zalloua, Zahi, *Being Posthuman: Ontologies of the Future*, London, Bloomsbury, 2021.

Notes

1. Jean Baudrillard, *Simulacra and Simulation*, translated by Sheila Glaser, Ann Arbor, University of Michigan Press, 1994, p. 1.

2. N. Katherine Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*, Chicago, University of Chicago Press, 1999, p. 8.

3. Mari Ruti, *Reinventing the Soul: Posthumanist Theory and Psychic Life*, New York, Other Press, 2006, p. 28.
4. Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford, Oxford University Press, 2014, p. 115-141.
5. N. Katherine Hayles, *op. cit.*, p. 42-43.
6. Nick Bostrom, *op. cit.*, p. 1.
7. Cary Wolfe, *What Is Posthumanism?*, Minneapolis & London, 2010, p. xxv.
8. Francesca Ferrando, *Philosophical Posthumanism*, preface by Rosi Braidotti, London & New York, Bloomsbury Academic, 2020, p. 67.
9. G. Douglas Barrett, *Experimenting the Human: Art, Music, and the Contemporary Posthuman*, Chicago, University of Chicago Press, 2023, p. 1.
10. Mark Kingwell, *Singular Creatures: Robots, Rights, and the Politics of Posthumanism*, Montreal, McGill-Queen's University Press, 2022, p. 1.
11. Fredric Jameson, *Archaeologies of the Future: The Desire Called Utopia and Other Science Fictions*, London & New York, Verso, 2005, p. 345.
12. Paul Matthews, *Transparent Minds in Science Fiction: An Introduction to Accounts of Alien, AI and Post-Human Consciousness*, Cambridge, Open Book Publishers, 2023, p. 58.
13. Emily Cox-Palmer-White, *The Biopolitics of Gender in Science Fiction: Feminism and Female Machines*, New York, Routledge, 2021, p. 3.
14. N. Katherine Hayles, *op. cit.*, p. 2-3.; Cary Wolfe, *op. cit.*, p. xxv.
15. N. Katherine Hayles, *op. cit.*, p. 345.
16. Paul Matthews, *op. cit.*, p. 27-28.
17. Zahi Zalloua, *Being Posthuman: Ontologies of the Future*, London & New York, Bloomsbury Academic, 2021, p. 21.
18. Fredric Jameson, *op. cit.* p. 286.
19. Donna J. Haraway, *Simians, Cyborgs, and Women: The Reinvention of Nature*, New York, Routledge, 1991, p. 243.
20. N. Katherine Hayles, *op. cit.*, p. 9.
21. Cary Wolfe, *op. cit.*, p. 136.
22. N. Katherine Hayles, *op. cit.*, p. 8.
23. Nick Bostrom, , *op. cit.*, p. 52-53.
24. N. Katherine Hayles, *op. cit.*, p. 22-23.
25. Nick Bostrom, *op. cit.*, p. 167.
26. Jean Baudrillard, *op. cit.*, p. 2-3.
27. Fredric Jameson, *op. cit.* p. 345.
28. N. Katherine Hayles, *op. cit.*, p. 6.
29. Cary Wolfe, *op. cit.*, p. xv.
30. Jean Baudrillard, *op. cit.*, p. 3.
31. G. Douglas Barrett, *Experimenting the Human: Art, Music, and the Contemporary Posthuman*, Chicago & London, The University of Chicago Press, 2021, p. 100.
32. N. Katherine Hayles, *op. cit.*, p. 12.
33. Francesca Ferrando, *op. cit.*, p. 85
34. Emmanuel Levinas, *Totality and Infinity: An Essay on Exteriority*, translated by Alphonso Lingis, Dordrecht, Boston & London, Kluwer Academic Publishers, 1991, p. 261.
35. Cary Wolfe, *op. cit.*, p. 88.
36. Francesca Ferrando, *op. cit.*, p. 98.
37. Jean Baudrillard, *op. cit.*, p. 1.
38. Fredric Jameson, *op. cit.* p. xi.
39. Jean-François Lyotard, *The Postmodern Condition: A Report on Knowledge*, translation from French by Geoff Bennington and Brian Massumi, foreword by Fredric Jameson, Oxford, Manchester University Press, 1984, p. xxiv.

40. Nick Bostrom, *op. cit.*, p. 134-135.
41. Ryan White, *The Hidden God: Pragmatism and Posthumanism in American Thought*, New York, Columbia University Press, 2015, p. 82.
42. Jean Baudrillard, *op. cit.*, p. 27.
43. N. Katherine Hayles, *op. cit.*, p. 8-10.
44. Francesca Ferrando, *op. cit.*, p. 65-67.
45. Rosi Braidotti, *The Posthuman*, Malden & Cambridge, Polity Press, 2013, p. 169.
46. N. Katherine Hayles, *op. cit.*, p. 11.
47. Paul Matthews, *op. cit.*, p. 103-104.