Rethinking Human-AI Hybrid Intelligence: Toward a New Theory

Abstract

This academic work argues that human-AI hybrid intelligence requires a new theoretical framework, as existing discussions remain limited to automation, tool use, or technological threats. These perspectives fail to fully capture the cultural and cognitive transformations that occur when humans interact with artificial intelligence.

In response, my thesis introduces Transsentientism as a dynamic theoretical framework for rethinking human-AI hybrid intelligence. This term describes a relational condition in which generative systems and machine outputs mediate human intuition, empathy, creativity, and ethical judgment.

Drawing on philosophical software studies, posthuman theory, phenomenology, and AI ethics, this thesis argues that hybrid intelligence acts as a bridge between the human and AI domains. The concept of Transsentientism is designed to remain open to revision, emphasising the need to understand intelligence as relational and increasingly hybrid.

Introduction

Artificial intelligence has become one of the most urgent cultural, ethical and intellectual questions of the contemporary period. It is no longer limited to laboratories, engineering departments, speculative fiction or specialist technical communities. AI systems now participate in writing, image generation, education, research, translation, design, search, administration, recommendation systems and decision-making. The rise of generative AI has made this transformation particularly visible because users can now interact with systems that produce text, images, code, audio and simulated dialogue in ways that appear fluent, responsive and contextually adaptive. However, the conceptual language used to describe this transformation often remains inadequate. AI is frequently framed either as a tool for productivity or as a threat to human labour, creativity and agency. Both perspectives are important, but neither is sufficient. They do not fully explain what happens when human thought and artificial cognition begin to operate together.

The thesis below argues that human-AI hybrid intelligence must be rethought because the relationship between humans and artificial systems can no longer be understood through the simple user-tool model. In conventional tool-use, the human subject appears to remain clearly separate from the instrument being used. A hammer, a camera or a word processor may extend human capacity. Still, it does not appear to

answer, suggest, simulate conversation or generate symbolic material in the same way as contemporary AI systems. Generative AI complicates this model because it does not merely execute commands in a mechanical sense; it produces language, images and patterns that enter into human processes of interpretation. When a user uses AI to brainstorm ideas, when a designer uses image generation to explore visual possibilities, or when a writer uses a language model as a dialogic partner, the process cannot be reduced to automation alone. Human intention and machine output form a shared field of production.

The title of this thesis, *Rethinking Human-AI Hybrid Intelligence: Towards a New Theory*, therefore, signals an argumentative shift. The aim is not simply to explore AI as a contemporary topic, but to challenge the categories through which AI and intelligence are commonly understood. Artificial intelligence is often discussed in terms of capability: what it can generate, calculate, predict or automate. Yet the more significant cultural question concerns transformation: how does AI alter the conditions under which humans think, create, interpret and make decisions? If AI systems increasingly participate in cognitive and creative processes, then intelligence itself can no longer be imagined only as an internal property of the individual human subject. It must also be understood as relational, mediated and distributed across humans, machines, interfaces, datasets and cultural contexts.

This is where the concept of *Transsentientism* enters the thesis. **Transsentientism** is proposed as a provisional theoretical framework for understanding the interaction between human sentience and artificial cognition. The term combines “trans”, suggesting movement across, beyond or through, with “sentient”, a term associated with perception, awareness and feeling. However, the concept must be defined carefully. Transsentientism does not claim that AI is sentient in a biological, emotional or human sense. It does not argue that machines possess lived experience, consciousness or moral responsibility. Instead, it asks how human sentience is transformed when it comes into contact with artificial systems capable of simulating reasoning, generating language, producing images, imitating empathy, and engaging in creative processes.

The condition for a new theory is also supported by work in software studies and posthuman theory. Manovich (2013) argues that software has become a cultural layer through which media are produced, edited and circulated. Frabetti (2015) similarly insists that software must be understood not only technically, but also philosophically and culturally. Bach’s work on synthetic intelligence adds another important dimension to this discussion. In *Principles of Synthetic Intelligence*, Bach (2009) approaches artificial cognition not as a simple simulation of human thought, but as an architecture of motivated, adaptive and goal-oriented processes. This is useful for rethinking AI not merely as a tool, but as a form of organised cognitive agency that increasingly interacts with human perception, decision-making and creativity. All these perspectives are important because AI is not simply a neutral tool added to culture from the outside. It is part of a software-based environment that shapes perception, production and

knowledge. Posthuman theory also challenges the idea of the human as an autonomous and self-contained subject. Hayles (1999) argues that the posthuman reconfigures the relationship between embodiment, information and technology, while Braidotti (2013) understands the posthuman subject as relational and embedded rather than isolated and universal. These arguments provide a foundation for understanding human-AI hybrid intelligence as a transformation of the human rather than simply an expansion of the machine.

At the same time, this thesis does not argue for the disappearance of the human as it is. On the contrary, it insists that human responsibility becomes more important within hybrid systems. AI may participate in production, but it does not bear ethical accountability. It can generate text, images or recommendations, but it cannot understand the social consequences of its outputs in the way humans and institutions can. For this reason, any theory of human-AI hybrid intelligence must include questions of agency, responsibility, transparency and judgement.

The present thesis develops the concept of a bridge layer between the human and AI domains. The human domain includes embodied experience, memory, emotion, cultural context, imagination and moral responsibility. The AI domain includes statistical prediction, pattern recognition, generative output, model architecture and computational processing. The bridge layer is the space where these domains meet. It includes prompting, interpretation, editing, evaluation, refusal, correction and responsibility. This bridge layer is where hybrid intelligence actually occurs. It is also where the human must remain critically active.

Transsentientism Academy, the digital media project connected to this thesis, translates this theoretical argument into a public-facing research hub. The website is not merely a place where the thesis is displayed. It is part of the project's method. Through its lexicon, visual frameworks, research notes and accessible design, the website turns an abstract theoretical problem into a navigable digital environment. This is important because a new theory requires not only academic explanation but also public communication. A concept such as Transsentientism must be readable, visualised and open to further development.

The ongoing thesis, therefore, extends in several stages. First, it examines the shift from artificial intelligence to hybrid intelligence. Second, it situates the argument within software studies, posthumanism and the cultural reconfiguration of the human. Third, it defines Transsentientism as an evolving theoretical proposal rather than a fixed doctrine. Fourth, it develops the bridge layer as a model for understanding human–AI interaction. Fifth, it examines ethics, agency and responsibility in hybrid systems. Finally, it reflects on the term's limits and the theory's future development.

Eventually, this thesis argues that rethinking human-AI hybrid intelligence is not only a technical matter. It is a cultural, philosophical and ethical necessity. As AI becomes more embedded in everyday life, humans need better concepts for understanding how

intelligence is changing. Transsentientism is offered as one possible beginning: not a final answer, but a theoretical opening.

1. From Artificial Intelligence to Hybrid Intelligence

The desire to create intelligence outside the human body is not new. Long before the emergence of modern computing, human cultures imagined artificial beings, animated statues, mechanical servants and nonhuman forms of agency. Myths such as Pygmalion and Galatea, or the figure of Prometheus, reveal a persistent fascination with the possibility of producing life, intelligence or animation through human craft. These myths are not simply decorative references. They show that artificial intelligence belongs to a longer cultural history in which humans repeatedly imagine the making of something that resembles, extends or challenges the human. AI, therefore, begins not only in engineering, but also in myth, imagination and symbolic fascination.

This historical background is important because it prevents artificial intelligence from being understood only as a recent technical phenomenon. Contemporary AI may depend on computation, data and machine learning, but the human desire behind it is much older. The dream of creating an artificial mind, an artificial companion, or an artificial helper has appeared across mythology, literature, philosophy and technology. In this sense, AI can be read as the technological continuation of an ancient cultural question: can intelligence, creativity, or agency exist outside the human biological body?

In the modern period, this question became increasingly connected to machinery, mathematics and computation. Charles Babbage's *Difference Engine* and *Analytical Engine* introduced the possibility of mechanical calculation. At the same time, Ada Lovelace's notes on the *Analytical Engine* suggested that such a machine might manipulate symbols as well as numbers (Menabrea and Lovelace, 1843). Lovelace's insight is significant because it implies that computation was never only about arithmetic. From its early conceptual formation, computing carried the possibility of symbolic transformation. It could be imagined not merely as calculation, but as a process capable of producing patterns, relations and forms of meaning.

Alan Turing's question, "Can machines think?", remains one of the most influential moments in the theoretical history of artificial intelligence (Turing, 1950). Turing did not answer this question by trying to define thought as an inner essence. Instead, he proposed the imitation game, shifting attention towards behaviour, language and interaction. This move is especially relevant to contemporary generative AI because many current systems create the impression of intelligence through linguistic performance. A machine does not need to possess human consciousness in order to affect how humans experience dialogue, authorship, authority or knowledge. Its output becomes culturally powerful because humans interpret it as meaningful.

The history of AI also includes early attempts to build systems capable of reasoning and acting in the world. Shakey (SRI International, n.d.), developed at the Stanford Research Institute in the late 1960s, is often remembered as one of the first mobile robots capable of reasoning about its actions, navigating and interacting with its environment. Shakey is important because it demonstrates that artificial intelligence was not imagined only as passive computation. It was also associated with planning, movement, perception and decision-making. However, the contemporary public encounter with AI has shifted significantly. Today, many users do not primarily encounter AI as a robot moving through physical space, but as a symbolic and generative system entering language, images, education, design and creative workflows.

This shift changes the terms of the debate and how we think about artificial intelligence. It is no longer just about whether machines can do intelligent things; it is more about how these artificial systems interact with us and help us find meaning in our lives. As Russell and Norvig point out, AI is about studying intelligent agents, but we also need to look at them from a cultural perspective (2022, pp. 983, 986).

At this point, it's important to distinguish between artificial intelligence and hybrid intelligence. Artificial intelligence is the technical field concerned with systems capable of performing tasks associated with intelligent behaviour. Hybrid intelligence, by contrast, denotes the relational condition in which human and artificial processes work together. In instances where individuals engage with image-generation technologies to explore visual craft, when a user employs AI to generate sounds or videos, or when a writer uses a language model as a dialogic partner, the result is not produced by the machine alone. Nor is it produced by the human in the traditional sense of isolated authorship. It emerges through prompting, responding, interpreting, selecting, correcting, and judging.

This process cannot be fully explained through the language of automation. Automation suggests that the intelligent machine takes over a task, often making it faster or more efficient. Hybrid intelligence suggests something more complex: the transformation of the task itself. Writing with AI is not simply faster writing. Designing with AI is not simply faster designing. Researching with AI is not simply faster searching. In each case, the human process is reorganised by the presence of a system that can generate suggestions, propose structures, simulate alternatives and influence decisions. The machine does not replace human thought in a simple way, but it changes the environment in which human thought takes place.

For this reason, human-AI hybrid intelligence requires a new conceptual vocabulary. The older notion of tool use is no longer sufficient, as generative AI does not behave like a passive instrument. At the same time, it would be misleading to describe AI as a human-like intelligence. It is AI-like intelligence. The challenge is to position the argument between these stances. AI is not merely a tool, but neither is it a human

subject. It is a computational system that enters human processes of interpretation and production. Hybrid intelligence names this in-between condition.

Transsentientism begins from this in-between condition. It argues that the most important issue is not whether AI is simply good or bad, conscious or unconscious, creative or derivative. These questions matter, but they do not exhaust the problem. The deeper issue is how human cognition, creativity and judgement are altered when artificial systems become part of the process through which humans think and make meaning. Rethinking human-AI hybrid intelligence, therefore, means moving from the question of machine capability to the question of relational transformation.

The movement from artificial intelligence to hybrid intelligence is therefore the first step towards a new theory. It allows AI to be analysed not only as a technical system but also as part of a cultural relation. It also allows the human to be understood not as an untouched subject standing outside technology, but as a being whose cognition is increasingly mediated by artificial systems. This does not mean that the human disappears. It means that human agency must be rethought under conditions of technological mediation. This is the theoretical space in which Transsentientism essentially evolves.

2. Software, Posthumanism and the Reconfiguration of the Human

To develop a new theory of human-AI hybrid intelligence, it is necessary to move beyond the idea that artificial intelligence is simply a technical tool. AI exists within a much wider software culture that shapes how contemporary life is organised, imagined and mediated. Software does not merely support cultural production in the background; it actively structures the conditions under which media are created, edited, circulated, and interpreted. Manovich (2013) argues that software has become a cultural layer that influences almost every form of contemporary media practice. This argument is important for the present thesis because AI is not external to culture. It is one of the most powerful current forms of software culture, and it increasingly participates in the production of text, images, knowledge, decisions and aesthetic forms.

From this perspective, artificial intelligence cannot be understood as a neutral instrument that obeys human commands. A generative AI system does not function like a hammer, pencil or camera in the traditional sense. It does not merely extend the hand or preserve an image. It generates symbolic material that enters human interpretive processes. It offers possible sentences, images, structures, associations and answers. These outputs are not independent thoughts, but they still influence human thinking. The user is therefore not only operating a tool; the user is engaging with a system that responds, suggests, imitates and reorganises the process of creation and thinking. This is why the model of simple tool-use is insufficient for understanding contemporary AI. How about the near-future Superintelligence?

Frabetti (2015), I highlight again, argues that software should be treated not only as a technical object but also as a philosophical and cultural form. This is especially relevant to AI because software increasingly shapes what can be seen, said, designed, searched for, remembered and produced. AI systems are built through code, datasets, models, interfaces, and platforms, but their effects cannot be limited to these technical components alone. They affect cultural imagination. They influence what users expect from knowledge, creativity and communication. They also reshape the relation between human agency and technical mediation. In this sense, AI becomes not only something humans use, but something through which humans increasingly think.

Braidotti (2013) also challenges the traditional humanist idea of a universal and independent human subject. Her account of the posthuman emphasises relationality, embodiment and embeddedness. The human is not an isolated centre of reason standing above the world, but a being formed through relations with other humans, nonhuman life, technologies, environments and systems. AI complicates this further because it introduces “non-human” systems that can simulate forms of reasoning and expression. These systems are not alive, embodied or morally responsible in the way humans are, but they still participate in the environments through which humans think and act. For Braidotti, posthuman interaction is not based on “the reactive bond of vulnerability”, but on “an affirmative bond that locates the subject in the flow of relations with multiple others” (Braidotti, 2013, p. 50).

The posthuman framework is useful because it allows this thesis to avoid two simplistic positions. The first is human exceptionalism, which assumes that human intelligence remains untouched by machines. The second is technological determinism, which assumes that technology develops according to its own inevitable logic and simply determines the future. Human-AI hybrid intelligence requires a more meticulous position. AI does not replace humans in a simple sense, but neither do they remain unchanged. Their practices of writing, designing, researching, deciding and imagining are increasingly mediated and improved (!) by artificial systems. The human is not erased, but **reconfigured**.

Haraway’s concept of the cyborg also provides an important reference point for thinking about hybridity. The cyborg challenges boundaries between human and machine, organism and technology, nature and culture (Haraway, 1991). Although Haraway’s cyborg emerged from a different historical and technological context, it remains useful because it shows that hybrid identity is not an exception but a condition of modern technological life. However, the hybrid intelligence discussed in this thesis is not only bodily or mechanical. It is cognitive, symbolic and noetic. It concerns the relation between human awareness and artificial cognition.

This is where Transsentientism begins to differ from earlier theories of human–machine hybridity. The cyborg helps to describe the fusion or entanglement of body and technology, but human-AI hybrid intelligence requires attention to interpretation, language, creativity and judgement. Generative AI does not necessarily attach itself to

the body. Instead, it enters the symbolic spaces through which humans produce meaning. It appears in the sentence, the image, the search result, the recommendation, the design draft and the research summary. It therefore transforms not only what humans can do, but how humans experience the process of thinking and making.

In the same vein, Baudrillard's theory of simulation is useful here because generative AI intensifies questions about representation, authenticity, and reality. For Baudrillard (1994), contemporary culture is shaped by signs and simulations that may become detached from stable referents. AI-generated texts, images and voices create similar tensions. They can appear meaningful, expressive or realistic without originating in human lived experience. A generated image may look like a photograph, a generated voice may sound like a person, and a generated text may resemble a human argument. This does not mean that all AI output is empty or false. Rather, it means that the relationship among origin, meaning, and authenticity becomes more difficult to determine. The work emerges through a **hybrid process**. Is this bad? Is this good?

At this stage of my thesis, I must bring to light Katherine Hayles' connection to my work. Hayles (2025) challenges anthropocentrism by arguing that the human-centred view of the world needs to be questioned and reworked. Anthropocentrism places human beings at the conceptual, political, social, and economic centre of existence, treating them as the most significant species on Earth and as the beings most entitled to use the planet's resources for their own purposes.

For Hayles, cognition is central to this problem. If anthropocentrism rests on the assumption that human cognition is superior or uniquely important, then any serious critique of anthropocentrism must also reconsider what cognition is and how human cognitive abilities should be understood in relation to other forms of life and intelligence.

In *Unthought: The Power of the Cognitive Nonconscious* (2017), Hayles argues that human thought is not limited to conscious reasoning. Beneath conscious awareness, nonconscious cognitive processes operate at a neurological level and are essential for consciousness itself to function. This means that human cognition is already more complex, layered, and distributed than traditional humanist models suggest.

Hayles, therefore, proposes a move away from the liberal Enlightenment view of the autonomous, rational human subject toward what she calls an **integrated cognitive framework** (ICF). This framework allows conscious human cognition to be examined alongside nonconscious forms of cognition within humans, across the wider living world, and within computational media. The aim is not to erase differences between these forms of cognition, but to create a way to compare their connections, distinctions, capacities, and origins.

The reconfiguration of the human through AI is thus both an opportunity and a risk. It creates new forms of creative extension, cognitive support and notional experimentation. But it also raises questions about agency, dependency, authorship, authenticity and responsibility. Transsentientism responds to this condition by proposing that human-AI hybrid intelligence should be understood as relational rather than individual, mediated rather than unaltered, and ethically demanding rather than automatically progressive. The human remains central, but not as a sole master of tools. The human becomes a responsible interpreter within a wider field of artificial intelligence cognition.

This section has argued that software studies and posthumanism provide the theoretical foundation for rethinking human-AI hybrid intelligence. Software studies demonstrate that AI is part of a cultural and technical infrastructure that shapes media and knowledge. Posthumanism suggests that the human experience has always been relational and technologically mediated. Together, these perspectives make it likely that AI is understood not simply as an external tool, but as part of a broader shift in how human intelligence is organised. The next step is to define Transsentientism more precisely as an evolving theoretical proposal within this transformation.

3. Transsentientism as an Evolving Theoretical Proposal

Transsentientism is introduced in this thesis as a provisional theoretical proposal for rethinking human-AI hybrid intelligence. It is not presented as a concluded doctrine, nor as a fixed academic category. Preferably, it functions as a working concept designed to open a space of inquiry around the changing relationship between human sentience and artificial cognition. The term brings together “trans”, suggesting movement across, beyond or through, and “sentient”, associated with perception, awareness, feeling and responsiveness. Its construction, therefore, points towards a transitional condition: a moment in which human awareness is increasingly shaped by artificial systems that can simulate aspects of cognition, language, creativity, and emotional response.

However, the concept must be carefully discerned from the claim that AI is sentient. Transsentientism does not argue that artificial intelligence possesses consciousness, feeling, selfhood or lived experience. It does not suggest that large language models (LLMs) or generative systems are alive, aware or morally responsible. Instead, the theory asks how human sentience is transformed when it interacts with systems that can imitate reasoning, generate symbolic material, produce images, respond conversationally and influence decision-making. The emphasis is therefore not the inner life of the machine, but the changing condition of the human in relation to machine intelligence.

The first theoretical movement we propose within Transsentientism is the claim that intelligence is **relational**. Intelligence should not be understood only as a property

enclosed within an individual human mind or located entirely inside a machine. It also emerges through relations between subjects, tools, languages, environments, institutions and systems. Human beings have always thought with external supports: writing, books, maps, machines, archives and digital media. AI intensifies this condition because it does not merely store or transmit human knowledge; it generates responses that can influence the very direction of thought itself. Hybrid intelligence is therefore not a simple combination of human and machine. It is a relational field in which both human and artificial intelligence processes shape the production of meaning.

The second conceptual movement is the assertion that human agency is mediated. In AI-assisted work, agency is not located in one simple place. The user writes a prompt, but the model generates possible responses. The interface structures what the user can ask and receive. The dataset influences the patterns available to the system. The platform determines the conditions of access, visibility and control. The human user then interprets, edits, accepts or rejects the output. Agency is therefore dispersed across a network of human and nonhuman elements. However, this does not mean that responsibility is equally distributed. The machine may participate in production, but ethical accountability remains with humans and institutions.

The third notional movement is the insistence that responsibility should remain central. Transsentientism is not a celebration of human-AI fusion. It does not argue that humans should surrender judgement to artificial systems. On the contrary, it argues that the more AI becomes involved in cognition and production, the more important human judgement becomes. If AI participates in writing, design, education or decision-making, then human users must become more aware of how outputs are produced, what assumptions they may contain, and how they should be evaluated. Hybrid intelligence, therefore, requires critical cognisance rather than passive dependence.

This is why Transsentientism should be understood as a theory of critical co-cognition rather than simple co-creation. Co-creation often emphasises production: human and machine working together to make something. Co-cognition goes further by focusing on the process of thinking itself. It asks how AI changes the way humans generate ideas, accept and organise knowledge, test possibilities, and evaluate meaning. In this sense, Transsentientism is concerned not only with what humans and AI produce together but also with how the human mind is affected by the presence of artificial cognition in the production process.

The evolving nature of the term is part of the theory itself. During the development of this project, Transsentientism emerged as an intuitive attempt to name the relation between human awareness and artificial cognition. However, later research revealed that the label “Transsentientism” was used to refer to a philosophy that “powers the game Evolve” without progressing in any way (Pete Sol, 2023, LinkedIn). This prior use does not invalidate the present thesis. Still, it requires clarification: in this project, Transsentientism is developed in a different theoretical direction, as a framework for

understanding human-AI hybrid intelligence rather than as a game-related philosophy. The term is not without difficulty. It is long, unfamiliar and potentially metaphysical. It may suggest to some readers that the theory makes claims about machine sentience, even though it does not. For this reason, the term must remain open to revision. It may eventually be shortened, replaced or absorbed into a broader theory of human-AI hybrid intelligence. This openness is not a weakness. It reflects the fact that the cultural condition under study is still developing.

Transsentientism sits at the intersection of software studies, posthumanism, phenomenology, and critical AI studies. It is philosophical because it asks how human awareness is transformed. It is cultural because it examines how AI changes the way we create meaning. It is ethical because it insists that responsibility cannot be delegated to the machine. The theory does not exist only as abstract writing; it is translated into a website, a lexicon, visual frameworks and accessible research content.

Transsentientism can therefore be understood as a bridge concept. It does not resolve all questions about AI, consciousness or creativity. Instead, it gives those questions a structure. It allows human-AI hybrid intelligence to be discussed without reducing AI to a passive tool or falsely elevating it to human consciousness. It names the unstable space between automation and agency, simulation and meaning, artificial cognition and human responsibility, which should inspire the audience to explore these complex issues. That unstable space is precisely where a new theory is needed.

4. The Bridge Layer: Human Domain, AI Domain and Hybrid Awareness

A new critical theory of human-AI hybrid intelligence requires a model that can describe relations without collapsing the difference between humans and machines. It is not enough to say that humans “use” AI, because this suggests a simple instrumental relationship. It is also not accurate to say that humans and AI “merge”, because this exaggerates the machine and weakens the distinction between biological consciousness and computational processing. What is needed is a more precise theoretical model that explains how human intention and artificial cognition interact while remaining independent of each other. For this reason, this thesis proposes the concept of the bridge layer.

The bridge layer describes the space of interaction between the human and AI domains. It is not a physical space, but a conceptual and practical area in which human intention, machine output and interpretive judgement meet. It includes the processes of prompting, reading, evaluating, editing, accepting, rejecting, correcting and redirecting AI-generated substance. The bridge layer is where hybrid intelligence actually takes place. It is the zone where artificial cognition becomes meaningful because it is interpreted by a human, and where human cognition is altered because it receives suggestions, structures or alternatives from an artificial system.

The bridge layer in this critical theory is the point at which the two domains interact. A prompt is not merely a command. It is a translation of human intention into a form that the machine can process. An AI response is not merely data. It becomes meaningful only when a human interprets it, evaluates it and decides what to do with it. Editing is not simply a technical correction. It is the return of human judgement into the process. Refusal is also important. The human user must be able to detect any hallucination, reject an output, question it, identify bias, recognise error, and resist the authority of fluency. Without this critical activity, hybrid intelligence risks becoming passive dependence. At this point, I used ChatGPT to reply to the question: "Analyse Transsentientism website and generate your understanding about the bridge layer concept: https://www.transsentientismacademy.co.uk/." The AI machine's reply is quite long; however, I will reproduce the generated conclusion below:

"The Bridge Layer is the symbolic, ethical and interpretive space where human intelligence and artificial intelligence meet, negotiate and reshape one another. It is the zone in which emotion encounters pattern, memory encounters data, narrative encounters prediction, and meaning is produced through human–AI interaction. Rather than treating AI as a passive tool or a replacement for the human, the Bridge Layer frames hybrid intelligence as a collaborative process that requires interpretation, responsibility and critical reflection."

Although the wording requires academic refinement, the core idea is useful.

Finally, the bridge layer is central to the new theory proposed in this thesis. It allows human-AI hybrid intelligence to be understood without reducing AI to a passive tool or falsely elevating it to human consciousness. It shows that the crucial site of intelligence is not the machine alone, nor the human alone, but the relation between them. In this relation, human judgement, ethical responsibility and critical awareness remain essential. Transsentientism names this condition and argues that it should be studied as one of the defining cultural transformations of the present.

5. Ethics, Agency and Responsibility in Hybrid Systems

A new theory of human-AI hybrid intelligence must be both conceptual and ethical. If AI systems increasingly participate in writing, design, education, research, decision-making, and communication, then the question is not only how humans and machines produce together, but also who remains responsible for the consequences of that production. Human-AI hybrid intelligence cannot be treated just as a creative opportunity or a technical innovation. It must also be understood as a sphere of ethical tension, because AI systems are embedded in social structures, institutional practices, economic interests, datasets and platform infrastructures. Artificial Intelligence does not enter human life as a neutral instrument. AI arrives already shaped by the conditions of its design, training, ownership, and deployment.

This is why the concept of agency becomes difficult in hybrid systems. In a traditional model of tool-use, agency appears to belong primarily to the human user. The tool extends human intention, but it does not appear to generate independent symbolic material. A pen does not suggest an argument; a camera does not decide the ethical meaning of an image; a word processor does not produce an interpretation. Generative AI complicates this model because it produces outputs that can influence the user's next thought, decision or creative direction. The user may begin with an intention, but the AI system can introduce unexpected language, structure, imagery or associations that the user might accept, disseminate, and become a shaping factor. It is also interesting to examine how Nolan's "Person of Interest" (2011–2016) explores the ethical and political implications of predictive artificial intelligence, surveillance, and machine agency, making it a useful cultural reference for a theory of human-AI hybrid intelligence.

The European Union's AI Act could be relevant here because it demonstrates that AI is increasingly understood in terms of risk, governance, and accountability (European Parliament and Council, 2024). Regulation is deemed necessary because AI systems do not operate only at the level of individual choice. Institutions, companies and states deploy them, and they can affect education, employment, healthcare, policing, media and public communication. However, legal regulation alone cannot fully explain the cultural and experiential dimensions of human-AI interaction. It can define obligations and risks, but it cannot fully describe what it feels like to think with a machine, to trust a generated response, or to become dependent on artificial feedback. This is where a cultural theory of hybrid intelligence remains necessary.

Human-AI hybrid intelligence requires a double approach. On the one hand, it needs governance, transparency, and institutional accountability. On the other side, it needs individual and cultural literacy. Users must learn how to work with AI critically, but institutions must also avoid placing the responsibility entirely on users. Platforms, universities, companies and governments shape the conditions of AI use. Responsibility must therefore be understood at multiple levels: personal, educational, institutional, technical and political.

Transsentientism must resist technological determinism. AI does not determine the future by itself, for now. Its social meaning depends on how it is designed, regulated, used, interpreted and resisted. At the same time, the theory must resist a naïve humanism that imagines humans remain untouched by technology. AI is already changing humans through new habits of writing, searching, learning, deciding, imagining, and relying on a machine's empathy. The ethical challenge is to recognise this transformation, but the human agent should not let AI replace judgement, imagination, or moral decision-making. Human agency means the ability to choose, question, refuse, interpret, and take responsibility.

6. Limits, Critique and Future Development

A new theory must acknowledge its own limits. This is especially important for a project that introduces an evolving concept. Transsentientism is useful because it opens a space for rethinking human-AI hybrid intelligence, but it is not a perfect or final term. It may sound long, abstract or too metaphysical, particularly because it contains the word “sentient”. This creates the risk that readers may misunderstand the theory as a claim about machine consciousness. However, the concept does not argue that AI is sentient. It argues that human sentience is increasingly transformed through interaction with artificial cognition.

This distinction must remain central to the theory. Transsentientism is not a theory of artificial consciousness. It is a theory of relational transformation. Its concern is not whether AI possesses experience, emotion or moral awareness, but how human awareness is altered when it engages with systems capable of producing language, images, suggestions, simulations and responses. Nevertheless, because the word itself may invite misinterpretation, the term should remain open to refinement. A future version of the theory may use a shorter or more precise term if that better communicates the argument.

This openness is not a weakness. On the contrary, it reflects the nature of the project itself. Human-AI relations are still developing, and the theoretical language used to describe them should not claim finality. A concept that studies transformation must also be able to transform itself. Transsentientism should therefore be understood as a theoretical prototype: an attempt to name and structure a cultural condition that is still emerging. Its value lies in beginning the conversation, not in closing it.

Another limitation concerns the scale of the topic. Human-AI hybrid intelligence includes technical, philosophical, ethical, political, economic, educational, psychological and environmental dimensions. A thesis of this length cannot address all of these equally. This project focuses mainly on conceptual framing, human agency, cultural interpretation and ethical responsibility. Future research could develop the political economy of AI, the environmental cost of computation, the labour involved in dataset production, the role of platform capitalism, or the unequal distribution of AI access and harm.

The question of inequality is particularly important. Not all users encounter AI from the same social position. Access to technology, language, education, disability, class, geography, age and digital literacy all shape how people experience artificial intelligence. A theory of human-AI hybrid intelligence must therefore avoid assuming a universal user. The experience of a student using AI for study, a designer using AI for creative work, a worker monitored by algorithmic systems, and a person excluded from digital access are not the same. Future development of Transsentientism must take these differences more seriously.

There are methodological limitations also. This thesis is primarily theoretical, reflective and argumentative. It does not include interviews, user testing or empirical research into how people experience AI-assisted thinking. Such research could strengthen the

project in the future. For example, Transsentientism Academy could include reflective case studies, student responses, creative experiments, interviews with AI users, or visual research into human-AI workflows. These additions would allow the theory to be tested against lived experience rather than developed only through conceptual analysis.

The future of the theory exposed here, therefore, depends on maintaining a balance between originality and caution. Transsentientism should be ambitious enough to propose a new way of thinking, yet careful enough not to overclaim. It should challenge narrow ideas of AI as a tool, while avoiding the mistake of treating AI as a human-like subject. It should defend human responsibility while recognising that human agency is increasingly mediated. It should welcome creative possibility while remaining critical of power, bias and extraction.

These limits do not undermine the argument of the present thesis. They clarify the next stage. The value of Transsentientism lies in its capacity to name a condition that is already shaping contemporary digital culture: intelligence is becoming relational, mediated and hybrid. The task now is to refine the language, deepen the research and develop the digital hub as a space where this theory can continue to progress.

Conclusion

This thesis argues that human-AI hybrid intelligence requires a new theoretical framework because current ways of discussing artificial intelligence are insufficient. AI is often described as a tool, an automation system, a simulation engine, or a technological threat. These categories remain useful, but they do not fully explain how human intelligence is transformed through sustained interaction with artificial cognition. The central issue is not only what AI can produce, calculate, or automate, but also how it changes the conditions under which humans think, write, design, decide, imagine, and interpret.

The thesis proposes Transsentientism as a provisional, evolving theoretical framework for understanding this transformation. The concept does not claim that AI is sentient, conscious or emotionally aware in a human sense. Instead, it argues that human sentience is increasingly mediated by artificial systems capable of generating language, images, suggestions, simulations and responses. Transsentientism, therefore, shifts attention from the question of whether machines are becoming human to a more immediate cultural question: how are humans changing as they think and create with machines?

“Rethinking Human-AI Hybrid Intelligence: Towards a New Theory” argues that the future of AI should not be understood solely in terms of machine progress. It must also be understood through the transformation of human thought, creativity and responsibility. The question is not whether humans or machines will dominate the future of intelligence. The more important question is how humans can remain

ethically, creatively and critically present within the hybrid systems they are now helping to create. Ultimately, it is about **co-evolution**.

Bibliography

Baudrillard, J. (1994) *Simulacra and Simulation*. Translated by S.F. Glaser. Ann Arbor: University of Michigan Press.

Bach, J. (2009) *Principles of synthetic intelligence: PSI: an architecture of motivated cognition*. Oxford: Oxford University Press. Available at: https://cognitive-ai.com/publications/assets/Draft-MicroPsi-JBach-07-03-30.pdf (Accessed: 4 February 2026).

Braidotti, R. (2013) *The Posthuman*. Cambridge: Polity Press.

European Parliament and Council (2024) *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence*. *Official Journal of the European Union*, L 2024/1689. Available at: https://artificialintelligenceact.eu/the-act/ (Accessed: 23 April 2026).

Frabetti, F. (2015) *Software Theory: A Cultural and Philosophical Study*. London: Rowman & Littlefield International.

Haraway, D. (1991) 'A cyborg manifesto: science, technology, and socialist-feminism in the late twentieth century', in *Simians, Cyborgs and Women: The Reinvention of Nature*. New York: Routledge, pp. 149–181. Available at: https://r2-vlebooks-com.roe.idm.oclc.org/Reader?ean=9780203873106 (Accessed: 9 March 2025).

Hayles, N.K. (2025) *Bacteria to AI: Human Futures with Our Nonhuman Symbionts*. Chicago: University of Chicago Press.

Hayles, N.K. (2017) *Unthought: The Power of the Cognitive Nonconscious*. Chicago: University of Chicago Press.

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

Manovich, L. (2013) *Software Takes Command*. New York: Bloomsbury Academic. Available at: http://dx.doi.org/10.5040/9781472544988.0005 (Accessed: 16 October 2025).

Nolan, J. (2011–2016) *Person of Interest*. Television series. New York: CBS. Available at: https://www.netflix.com/search?q=person%20of&jbv=70197042 (Accessed: 27 February 2026)

Russell, S. and Norvig, P. (2022) *Artificial Intelligence: A Modern Approach*. 4th edn. Harlow: Pearson.

Science Museum (2015) *Ada Lovelace's remarkable story to be celebrated at the Science Museum*. Available at: https://www.sciencemuseum.org.uk/about-us/press-office/ada-lovelaces-remarkable-story-be-celebrated-science-museum (Accessed: 25 January 2026).

Science Museum (2023) *Charles Babbage's Difference Engines and the Science Museum*. Available at: https://www.sciencemuseum.org.uk/objects-and-stories/charles-babbages-difference-engines-and-science-museum (Accessed: 24 January 2026).

SRI International (n.d.) *Shakey the Robot*. Available at: https://www.sri.com/hoi/shakey-the-robot/ (Accessed: 10 January 2026).

Turing, A.M. (1950) 'Computing machinery and intelligence', *Mind*, 59(236), pp. 433–460. Available at: https://courses.cs.umbc.edu/471/papers/turing.pdf (Accessed: 29 January 2026).

Resource list

OpenAI (2026) *ChatGPT response to Marinela Serban, 4 May*. ChatGPT. Available at: Bridge Layer Concept Analysis (Accessed: 4 May 2026).

Sol, P. (2023) *The philosophy of Transsentientism in Evolve* [LinkedIn post]. Available at: LinkedIn, https://www.linkedin.com/posts/petersoll_the-philosophy-of-transsentientism-in-evolve-activity-7093199464386945024-g-c0 (Accessed: 16 October 2026).