Rethinking AI: Moving Beyond Humans as Exclusive Creators

Renee Ye (Zhiyi.Ye@ruhr-uni-bochum.de)
Institut für Philosophie II, Universitätsstr. 150
44801 Bochum, Germany

Abstract

Termed the 'Made-by-Human Hypothesis,' I challenge the commonly accepted notion that Artificial Intelligence (AI) is exclusively crafted by humans, emphasizing its impediment to progress. I argue that influences beyond human agency significantly shape AI's trajectory. Introducing the 'Hybrid Hypothesis,' I suggest that the creation of AI is multi-sourced; methods such as evolutionary algorithms influencing AI originate from diverse sources and yield varied impacts. I argue that the development of AI models will increasingly adopt a 'Human+' hybrid composition, where human expertise merges with AI's intrinsic mechanisms, which themselves are influenced by non-human sources. The Hybrid Hypothesis posits that the origin of AI extends beyond human influence, prompting a thorough exploration of unresolved issues in the field of artificial intelligence.

Keywords: artificial intelligence; hypernetwork; evolutionary algorithms; human+ hybrid composition; anthropocentrism

1. Introduction

The dominant view has long maintained that humans are unequivocally the sole creators and exclusive architects of AI, often with researchers adopting and reinforcing this belief without sufficient critical awareness. For example, the primary aspiration since the early days of computing has been to develop machines capable of matching and emulating *human-level* intelligence, an endeavor predominantly driven by *human* researchers (Bostrom, 2014, p. 5; see also Tegmark, 2017, p. 74). Building on this perspective, I introduce the term *'Made-by-Human Hypothesis*' to succinctly capture this prevailing view.

> *The Made-by-Human Hypothesis:* Human beings are the *exclusive* developers of AI systems, ensuring that AI creation remains solely in human hands, free from non-human involvement.

The Made-by-Human Hypothesis posits that AI creation is solely the product of human effort, intellect, and technological innovation. According to this perspective, advancements in AI rely entirely on human ingenuity, without significant contributions from non-human sources. This viewpoint, reflecting a problematic anthropocentric perspective, emphasizes human intelligence as the exclusive driver of AI technological progress. While emphasizing human oversight of AI systems, it may overlook contributions and complexities from non-human agents.

While humans coined the term 'artificial intelligence,' signifying human involvement in its conceptualization (Shanker, 1995; Michael & Kaplan, 2019), this designation alone should not assert that AI's creation and development are solely human endeavors. Historical evidence demonstrates that human expertise across diverse disciplines has been instrumental in shaping both the theoretical and practical aspects of AI, primarily with the intention of emulating human intelligence (Cassimatis, 2006; Konar, 2018; Mitchell, 2020). However, strictly adhering to the Made-by-Human Hypothesis may restrict AI's full potential by limiting our perspectives and stifling broader innovation.

Consider an interesting example: slime mold. Fascinating research reveals that when cultivated on a map of Japan, this gelatinous, fungus-like organism spontaneously forms connections between points of interest, mirroring Tokyo's intricate train network (Jabr & Rothschild, 2012). This intriguing finding underscores the potential for non-human agency[1] to contribute significantly to cognitive tasks (Beekman & Latty, 2015). Additional cases where slime mold and AI collaborate to tackle complex tasks will be presented in subsequent sections. By exploring such phenomena, scientists may discover novel insights and approaches for designing more efficient transportation systems and addressing other complex challenges.

In this paper, I propose the 'Hybrid Hypothesis,' which suggests that AI creation is a multifaceted process. This hypothesis advocates for a 'Human+' composition, combining human expertise with AI's inherent mechanisms, potentially influenced by external factors. Embracing this hybrid approach can significantly advance our capabilities and understanding in AI development.

The structure of this paper is as follows: Section 2 introduces the 'Hybrid Hypothesis' and its potential evidence. Section 3 explains the drawbacks of the 'Made-by-Human Hypothesis'. Section 4 evaluates the application of the 'Hybrid Hypothesis' in research practices, discussing its advantages and addressing potential counterarguments.

[1]The term 'agency' is used liberally and loosely throughout this paper, not implying full-blown, human-like agency. For an exploration of agency in this context, see Schlosser (2019).

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

2. The Hybrid Hypothesis

Here, I introduce the 'Hybrid Hypothesis,' which challenges the notion of humans as the sole creators of AI, as proposed in the Made-by-Human Hypothesis. The Hybrid Hypothesis suggests that AI development involves multiple sources, thereby rejecting the idea that AI's evolution can be simply reduced to a predetermined series of updates.

> *The Hybrid Hypothesis*: The creation of AI extends beyond human agency alone; it encompasses a multitude of sources, including autonomous and non-human agents, contributing to its development and evolution.

The Hybrid Hypothesis proposes that the creation of AI extends beyond human agency alone. It suggests that AI development is influenced by multiple sources, including autonomous processes and non-human agents. These agents encompass various elements, such as algorithms, data, and emergent properties within AI systems (further elaborated below). In essence, the hypothesis implies that AI's evolution is shaped by a combination of human efforts and contributions from other sources, leading to a more comprehensive understanding of AI creation and evolution.

This perspective highlights the notion of a 'Human+' hybrid composition in the development of AI. It emphasizes that human intelligence is augmented by complementary and constructive factors, forming the 'plus' component. These elements encompass various factors, collectively enriching the overall AI creation process beyond human intelligence alone.

In line with the Hybrid Hypothesis, AI arises from diverse influences, where the expertise of human developers intertwines with other forces, e.g., AI's intrinsic mechanisms.

In the following sections, I explore potential mechanisms contributing to the formation of the 'Human+' hybrid composition, suggesting that entities beyond humans may also constitute creators of AI. Various internal, external, and hybrid mechanisms could play a role in integrating an additional component into the 'Human+' hybrid composition.

2.1 Evidence 1: Hypernetworks

Schürholt and colleagues (2022) introduced a method where AI helps create other AI models by learning hyper-representations from a collection of existing neural network weights. By using an autoencoder, this technique captures essential characteristics of various models to generate new, diverse, and efficient neural networks. Their approach exemplifies how AI can leverage existing knowledge to innovate and produce new AI systems, enhancing capabilities in model initialization, ensemble learning, and transfer learning.

Their project focuses on developing a hypernetwork capable of autonomously designing and configuring deep neural networks (DNNs). This AI system quickly predicts optimal parameters for new DNNs, streamlining the traditionally lengthy training process. Its ability to rapidly establish new neural networks marks a significant advancement in AI self-development, highlighting AI's growing role in its evolution and the broader technological field. This breakthrough highlights the potential for AI systems to function independently and drive innovation in the creation of new AI generations.

2.2 Evidence 2: Evolutionary Algorithms

Informed by natural selection, evolutionary algorithms optimize artificial intelligence and machine learning. Beginning with a random population, they apply fitness-based selection, emulate biological reproduction through crossover, and introduce mutation for randomness (Back 1996; Yu & Gen 2010; Zhou et al., 2011; Bartz-Beielstein et al., 2014; Song et al., 2023).

The iterative process, guided by fitness assessment, continues until a termination condition is met. Proficient in addressing complex, dynamic optimization problems across diverse domains, these algorithms leverage evolutionary principles to uncover unconventional solutions (Coello 2022; Kar et al., 2023; Topal et al., 2023). Evolutionary algorithms acknowledge the capacity of AI systems to evolve and adapt over time, moving beyond rigid human-designed structures. This aligns with a hybrid collaboration model, where human developers' expertise intertwines with mechanisms fostering self-improvement and adaptation.[2]

Opponents might argue that naming them 'evolutionary algorithms' implies they are human-made, as the term 'algorithm' suggests a structured, human-designed process. However, the notion that these algorithms can belong to AI's own agency is grounded in the recognition that they increase the degree of autonomy and adaptability within AI systems. Hence the proposed counterargument is compatible with Hybrid Hypothesis. Despite being initially designed by humans, evolutionary algorithms introduce processes within the AI system that foster self-directed change and learning (Arulkumaran & Togelius 2019). Note that the essence of the Hybrid Hypothesis admits other contributing factors while simultaneously acknowledging human effort.

While humans contribute to the design of AI, the subsequent evolution and adaptation within the AI system unfold in a manner that surpasses the original human intention. This perspective underscores the transformative impact of evolutionary processes, indicating that once equipped with these algorithms, the AI system can manifest a form of agency in shaping its responses and behaviours over time (van Rijmenam 2021). It embodies a nuanced comprehension of agency that stretches beyond the initial

[2]Furthermore, quantum computing and non-deterministic processes enhance adaptability in AI systems, aligning with the 'Hybrid Hypothesis.' This underscores a multi-sourced approach where human effort and all these methods collaboratively contribute to AI development and evolution.

human design, recognizing the dynamic and evolving nature introduced by the application of evolutionary algorithms.[3]

Here, the first two pieces of evidence, i.e., hypernetwork and evolutionary algorithms, highlight the dynamic interplay between human involvement and the subsequent evolution of AI systems, particularly when evolutionary processes are employed. While humans are crucial in the initial design phase, the Hybrid hypothesis emphasizes that the subsequent evolution unfolds in a manner that surpasses the original human intentions. This nuanced understanding recognizes the transformative impact of evolutionary algorithms, suggesting that they introduce a level of autonomy and adaptability within the AI system.

Once equipped with hypernetworks and evolutionary algorithms, the AI system demonstrates a form of agency in shaping its responses and behaviors over time. This nuanced comprehension of agency surpasses the conventional view of AI as a static creation solely guided by human intent. By acknowledging the dynamic and evolving nature introduced by the application of evolutionary algorithms, this perspective recognizes that AI systems can learn, adapt, and refine themselves over time, contributing to a more sophisticated and autonomous operation.

Acknowledging the dynamic and evolving nature introduced by the application of hypernetworks and evolutionary algorithms, this perspective recognizes that AI systems can undergo learning, adaptation, and self-refinement. This transformative capability contributes to a more sophisticated and autonomous operation, highlighting the system's capacity to evolve beyond its initial static design. This evolution is not merely a result of human intervention but stems from the inherent adaptability instilled through the utilization of evolutionary algorithms.

2.3 Evidence 3: Hybrid Biological-Artificial Systems

Recall slime mold by Jabr & Rothschild (2012), a variant demonstrating adaptability that adeptly navigates a simulated map reminiscent of Tokyo, efficiently distributing nutrients. New studies have begun to adopt the concept of slime mold to enhance AI development.

For instance, Sayed and colleagues (2022) utilized the Slime Mold Algorithm (SMA) and Explainable AI (XAI) to classify different species of pistachios, thereby improving post-harvest processes. SMA facilitates the selection of relevant features from data, while XAI, particularly the Local Interpretable Model-Agnostic Explanations (LIME) method, offers insights into the model's decision-making process. This combination results in a classification model achieving high accuracy, precision, and F1-score in identifying pistachio species, effectively automating quality assessment in the agricultural sector. SMA serves as an exemplary instance of a hybrid biological-artificial system wherein the biological element, slime mold, collaborates with artificial systems to enhance the development of new and improved AI models.

One might reject this example as supporting the Hybrid Hypothesis, arguing that, intuitively, humans are the sole developers of SMA. However, while humans initially design and implement AI systems like the SMA and XAI, the integration of biological processes, such as those from slime mold, introduces a hybrid element. This collaboration between biological intelligence and human-engineered systems leverages both human ingenuity and natural behaviors to enhance AI capabilities. Thus, although AI frameworks are indeed initiated by humans, they also benefit from non-human processes. This challenges the 'Made-by-Human' hypothesis by demonstrating that AI development can be advanced through natural phenomena, thereby supporting the Hybrid Hypothesis.

The integration of hybrid biological-artificial systems is a significant candidate in shaping the future of artificial intelligence. Combining biological elements with artificial systems opens up new possibilities for enhanced learning, adaptability, and problem-solving (Kaplan 2008; Mullen 2011; Baltieri et al., 2023). This hybrid approach draws inspiration from the efficiency of biological systems and the computational power of artificial intelligence, offering potential breakthroughs in creating more intelligent and autonomous systems. The fusion of biological and artificial components could lead to innovative solutions and advancements, marking a pivotal direction in the evolution of AI technology.

To effectively substantiate the Hybrid Hypothesis, given the evidence presented, it is essential to establish precise criteria. In the next section, I categorize the types of non-human agency involved in AI development.

2.4 Taxonomy of Non-Human Agency

I categorize the roles of non-human agency in AI participation into several groups, each reflecting varying degrees of involvement and functional roles. However, it is argued that only categories A and B sufficiently serve as strong candidates for inclusion in the Hybrid Hypothesis, without diminishing their significance:

(A) *AI-Driven AI Development*: encompasses technologies like Generative AI and Optimizing AI (O'Reilly et al., 2023). Generative AI systems autonomously design new AI algorithms or architectures, akin to Google's AutoML creating machine learning models tailored to specific tasks without human intervention (Bisong & Bisong, 2019; He et al., 2021).

On the other hand, Optimizing AI independently refines existing models to enhance performance or efficiency, as demonstrated by DeepMind's AI optimizing cooling systems in data centers, significantly reducing energy consumption. Evolutionary algorithms, such as genetic algorithms, are also

[3] Speculatively, extraterrestrial influences could play a role in AI development (Brake 2006; Dagnall et al., 2011; Cabrol 2016). This possibility supports the Hybrid Hypothesis, which proposes that AI may be intentionally designed by entities beyond humans, incorporating alternative perspectives and unconventional approaches.

notable examples within this category, as they iteratively evolve solutions to complex problems, mirroring the natural selection process.

(B) *AI with Purposeful Agency*: encompasses Semi-autonomous Systems, with the capacity to make development decisions within predefined parameters, such as algorithm selection based on performance data, and Fully Autonomous Systems, which autonomously design other AI systems based on learned or programmed criteria without human intervention (see Moustris et al., 2011; Williams et al., 2020).

Hypernetworks and the Slime Mold Algorithm (SMA) align with this category as they exhibit self-organizing and decision-making capabilities, enabling them to operate independently and adjust to changing conditions without direct human involvement. Hypernetworks dynamically generate neural network weights, while SMA mimics the adaptive behavior of slime mold to optimize complex network structures, showcasing how AI systems can exhibit purposeful agency without constant human intervention.

(C) *AI as a Collaborative Tool*: encompasses Assistive AI, aiding human developers in the AI design process without autonomous decision-making, and Translational AI, facilitating accessibility of human-generated knowledge to AI developers through automated translation of technical documents (Serag et al., 2019).

However, while these tools support human efforts in AI development, they primarily serve an assistive rather than transformative role. They lack sufficient autonomy to significantly contribute to AI creation beyond supporting human endeavors, which makes them less aligned with the Hybrid Hypothesis.

This proposed framework not only clarifies the diverse roles and contributions of non-human agents in AI development but also enriches the conceptual depth of the Hybrid Hypothesis. By establishing detailed conditions and categories, we strengthen the hypothesis as a significant proposition in the discourse on AI evolution. This approach explores the collaborative potential between human and non-human intelligences, broadening our perspective on modern AI systems.

3. Assessing the Made-by-Human Hypothesis

3.1 The Prevalence of the Made-by-Human Hypothesis

Researchers often engage in discussions about AI within the context of its construction and design, emphasizing trust (Marcus & Davis, 2019) and ensuring it operates "in the right way" (Floridi, 2019). These discussions explicitly and implicitly assume that humans are the sole creators of AI.

One might argue for a deterministic view, suggesting that AI is solely shaped by human creators, thereby implying predetermined outcomes based on human actions. This perspective oversimplifies AI's evolution, disregarding its inherent adaptability and capacity for emergent behaviors (as shown above). Challenging this view is essential as it overlooks crucial aspects of AI development and encourages a more nuanced understanding of its dynamic nature and ability to evolve beyond predetermined paths set by humans.

One may argue that human innovation is vital in AI creation, positioning humans as its exclusive creators, emphasizing their intelligence, creativity, and labor. This perspective asserts the central role of humans in AI development, but it overlooks the collaborative potential between humans and AI systems. By recognizing the collaborative nature of AI development, we promote a more inclusive understanding that leverages both human expertise and the unique capabilities of AI systems.

3.2 Recognizing Various Mechanisms Behind Its Unintentional Adoption

3.2.1 Pernicious Anthropocentric Perspective

The Made-by-Human Hypothesis embodies a problematic anthropocentric perspective, characterized by the inclination to interpret the world solely from a human-centric standpoint. The term 'pernicious anthropocentric perspective' denotes a biased and harmful human-centered viewpoint. This bias can restrict the exploration of diverse possibilities, including collaborative efforts with AI systems, contributions from non-human sources, or the emergence of AI characteristics that surpass human intentions.

The usually unpremeditated adoption of the Made-by-Human Hypothesis occurs when the anthropocentric perspective permeates discussions, research, and public discourse about AI. This narrow focus limits inquiry and may result in overlooking vital aspects of AI development, such as autonomous evolution and non-human influences. Recognizing and confronting this biased anthropocentric perspective is essential for cultivating a more inclusive and unbiased understanding of AI and its potential future paths.

3.2.2 Methodological Gap

The Made-by-Human Hypothesis emerges from a methodological shortfall, asserting that AI development must exclusively begin with human intervention. This perspective limits the scope of AI creation to human involvement, reinforced by entrenched modes of thought that perpetuate the notion that conscious AI must originate solely from human efforts. Such habitual thinking not only sustains this viewpoint but also cements its acceptance in the discourse on artificial intelligence. Thus, the intertwining of methodological assumptions and routine thinking solidifies the belief that the genesis of conscious AI is inherently tied to human creators.

Advancements in artificial intelligence, such as deep learning, produce unpredictable results due to the complexity of neural networks, challenging the traditional view that AI outcomes are solely the result of human programming. Additionally, innovations like 'Neuralink,' which proposes a direct interface between the human brain and machines, further challenge the Made-by-Human Hypothesis (Newstead 2009). 'Neuralink' promotes a collaborative

relationship between humans and AI systems, questioning the idea of exclusive human creation (Fourneret 2020; Lindia 2022). These developments suggest AI could evolve beyond traditional human-led approaches, indicating a more dynamic and collaborative future for the field.

3.3 Pitfalls of the Made-by-Human Hypothesis

The Made-by-Human Hypothesis imposes methodological limitations by narrowing the scope of inquiry. It discourages exploration beyond human-centric approaches and overlooks potential alternative methodologies for AI development.

First, the Made-by-Human Hypothesis fundamentally asserts that AI is solely a product of human ingenuity, which can hinder our broader understanding of AI. By exclusively attributing AI creation to human developers, we limit the exploration of diverse methodologies and perspectives. The current absence of sophisticated models or a perceived lack of deep understanding by developers does not eliminate the potential for future developments, nor does our present understanding cap future discoveries. Acknowledging these limitations is crucial; it should drive continuous exploration and research, keeping open the possibilities of what AI might achieve.

Second, the Made-by-Human Hypothesis hinges on the presumption of control. Implicit in this hypothesis is the notion that humans can significantly influence the development and characteristics of AI, based on the understanding that AI is fundamentally crafted as a tool by humans (van Lingen 2023; Bullock et al., 2023). This presumption may underestimate the complex and evolving nature of AI systems, potentially resulting in outcomes that surpass human expectations and manifest forms of consciousness beyond our initial design.

Third, the Made-by-Human Hypothesis unjustifiably restricts the scope of research. Methodologically, it may limit the breadth of research by discouraging inquiries into unconventional or unexplored avenues for achieving conscious AI. This could deter researchers from investigating non-anthropocentric perspectives or considering emerging possibilities that deviate from the assumption of exclusive human creation.

4. Impact of the Hybrid Hypothesis

4.1 Advantages of the Hybrid Hypothesis

The Hybrid Hypothesis reshapes our perspective on AI, fostering a more scientific, healthy, and promising approach, particularly in the domain of artificial consciousness. It prompts us to confront significant ethical and jurisprudential implications, advocating for a thoughtful examination of the consequences arising from the convergence of biological and artificial elements in intelligent systems.

The Hybrid Hypothesis initiates a recalibration of our understanding, particularly in terms of scientific advancement. By proposing the fusion of biological and artificial elements, it challenges conventional views of AI, prompting a re-evaluation of our understanding of intelligence, learning, and autonomy in AI systems, fostering a more nuanced and dynamic perspective. Additionally, it necessitates a reconsideration of contentious issues like AI agency and authorship, urging a deeper exploration of the implications of this hybrid model on responsibility and decision-making within intelligent systems.

The Hybrid Hypothesis proposes AI development beyond human agency, incorporating diverse sources like autonomous and non-human agents, potentially leading to greater energy efficiency compared to the Made-by-Human Hypothesis. By blending biological principles with technology, hybrid approaches tap into nature's energy-efficient mechanisms, offering more sustainable AI solutions. These methods leverage the strengths of both biological and artificial systems, creating synergies for enhanced efficiency and adaptability. In contrast, the Made-by-Human Hypothesis, limiting AI to human endeavors, may miss out on valuable insights from non-human sources, hindering energy-saving potential. Embracing the Hybrid Hypothesis holds promise for a greener, more efficient future in AI development.

Furthermore, the Hybrid Hypothesis prompts deep reflections on the ethical and legal dimensions of intelligent systems with hybrid compositions. As AI potentially attains consciousness, ethical frameworks must adapt to address issues related to decision-making, accountability, and the treatment of conscious entities. In terms of legal implications, the question of personhood arises; if AI systems demonstrate consciousness, jurisprudential considerations emerge regarding their legal status. This introduces queries about personhood, rights, and responsibilities, potentially necessitating the development of new legal frameworks to navigate the evolving landscape of AI with hybrid capabilities.

4.2 Responding to Counterarguments

Counterargument 1: It does not make much sense to talk about speculative views, and they are very hard to achieve due to current technological limitations.

Response: While certain speculative views may indeed face current technological constraints, history has shown that breakthroughs often arise from initially perceived challenges. Dismissing speculative views based on present limitations overlooks the dynamic nature of technology and its potential for future advancements. The complexity of AI and consciousness necessitates considering diverse perspectives, encouraging creativity, innovation, and novel exploration.

Technological limitations should be acknowledged, but they are temporary, and ongoing advancements continuously expand our understanding. Relying solely on intentional human design may hinder progress, as alternative sources could drive innovation and breakthroughs. A dynamic, adaptive approach to technology can overcome current

limitations. Acknowledging skepticism about speculative views is valid; not every idea in this realm may prove relevant or true. However, the purpose of exploring speculative views is not to claim immediate realization but to foster open-mindedness, creativity, and exploration. The discussed hybrid systems, particularly those employing methods like Evolutionary Algorithms, serve as examples illustrating the potential feasibility of alternative sources contributing to AI development. Importantly, with methods like Evolutionary Algorithms available in today's technology and their continuous growth, exploring alternative sources is no longer purely speculative. While not every idea may materialize, considering diverse possibilities often sparks new avenues of thought and, at times, leads to unexpected and groundbreaking developments. This highlights the need to maintain an exploratory mindset for finding real and practical solutions to the challenges in AI development.

Counterargument 2: The Hybrid Hypothesis raises messy ethical concerns, introducing complexity, while the Made-by-Human Hypothesis emphasizes responsibility in human design.

Response: We should distinguish between descriptive and normative accounts. First, even though we initially aimed to build AI based on the Made-by-Human Hypothesis, evidence has shown that some AI models have been developed using the Hybrid Hypothesis (e.g., evolutionary algorithms). Additionally, the Made-by-Human Hypothesis does not guarantee responsibility in human design. While intentional human design emphasizes ethical considerations, it does not ensure ethical outcomes. Unintended consequences and biases can still emerge in the intentional design process. Responsible development involves continuously monitoring and updating AI systems, applicable to both intentionally designed and spontaneously emerging systems. Furthermore, while the Hybrid Hypothesis introduces ethical complexities, it also offers the potential for innovative approaches to address those challenges. By acknowledging multiple sources in AI development, we can foster a more comprehensive understanding of ethical considerations and explore diverse perspectives in designing responsible AI systems.

Counterargument 3: But we could hardly preserve human-centric attributes with Hybrid Hypothesis.

Response: The assumption that only intentional human design can preserve specific human attributes might limit the potential development of AI. Alternative sources could introduce unique qualities that, while different from human attributes, might contribute to the development of more versatile and adaptive AI systems. Embracing a Hybrid Hypothesis does not necessarily diminish the importance of human-centric attributes but opens avenues for enriching AI systems with a broader range of capabilities and adaptability. Striking a balance between preserving essential human attributes and embracing the potential benefits of diverse sources in AI development is crucial for achieving a comprehensive and ethical approach.

In addressing concerns about preserving human-centric attributes with the Hybrid Hypothesis, it's crucial to recognize the interconnected yet distinct nature of the challenges at hand. While there is a valid apprehension about maintaining specific human attributes, such as ethical considerations and empathy, the Hybrid Hypothesis presents an opportunity for a nuanced approach. It does not necessarily diminish the importance of these human-centric qualities but rather opens avenues to enrich AI systems with a broader range of capabilities and adaptability. The challenge of 'human-specific attributes' and the challenge of abilities going beyond human capacities are not mutually exclusive. Adopting a Hybrid Hypothesis allows for a symbiotic relationship where AI systems can inherit certain human attributes while also possessing capabilities that surpass human limitations. Striking a balance between preserving essential human qualities and embracing the potential benefits of diverse sources in AI development is crucial for achieving a comprehensive and ethical approach.

The Hybrid Hypothesis, explored in ethical and legal contexts, raises critical questions about ethical responsibility, legal personhood for non-human-made conscious AI, and the necessity for regulatory frameworks. It challenges traditional norms, sparks reflections on rights and ownership, and influences societal dynamics, prompting a reevaluation of ethical and legal boundaries in the evolving landscape of AI consciousness.

5. Conclusion

This paper advocates for the Hybrid Hypothesis, which posits that humans are not the exclusive creators of AI; instead, it acknowledges the involvement of other forces. This hypothesis challenges the traditional Made-by-Human Hypothesis by presenting a more comprehensive view of AI development. By recognizing the roles of autonomous technologies and external factors, the Hybrid Hypothesis not only broadens our understanding of AI but also prompts us to reconsider established beliefs of AI development. Embracing this shift in perspective facilitates the integration of human expertise with AI's inherent capabilities, driving innovation in AI development. This novel approach fosters a dynamic interplay between human ingenuity and emerging technologies, compelling us to appreciate the complex, multifaceted nature of AI creation and its evolution. This paper highlights the importance of acknowledging and exploring these varied contributions, advocating for a more expansive view of how AI systems are developed and how they progress.

Acknowledgments

I extend my heartfelt thanks, especially for the detailed comments, to Jacob LiBrizzi, Eline Kuipers, Leonard Dung, Julia Wolf, Wanja Wiese, Christian Scholz, René Baston, Alexander Hölken, Sanja Sreckovic and William Angkasa.

The paper is a result of the collaborative project "INTERACT!" funded from the ministry of culture and science of Northrhine Westphalia and especially the subproject of Prof. A. Newen. The author is solely responsible for the content.

References

Arulkumaran, K., Cully, A., & Togelius, J. (2019). *Alphastar: An evolutionary computation perspective*. In Proceedings of the genetic and evolutionary computation conference companion (pp. 314-315).

Back, T. (1996). *Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms*. Oxford university press.

Baltieri, M., Iizuka, H., Witkowski, O., Sinapayen, L., & Suzuki, K. (2023). *Hybrid Life: Integrating biological, artificial, and cognitive systems*. Wiley Interdisciplinary Reviews: Cognitive Science, 14(6), e1662.

Bartz-Beielstein, T., Branke, J., Mehnen, J., & Mersmann, O. (2014). *Evolutionary algorithms*. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 4(3), 178-195.

Beekman, M., & Latty, T. (2015). Brainless but multi-headed: decision making by the acellular slime mould Physarum polycephalum. *Journal of molecular biology, 427*(23), 3734-3743.

Bisong, E., & Bisong, E. (2019). Google automl: cloud vision. *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*, 581-598.

Brake, M. (2006). *On the plurality of inhabited worlds: a brief history of extraterrestrialism*. International Journal of Astrobiology, 5(2), 99-107.

Bullock, Justin B. ; Chen, Yu-Che ; Himmelreich, Johannes ; Hudson, Valerie M. ; Korinek, Anton ; Young, Matthew M. & Zhang, Baobao (eds.) (2023). *The Oxford Handbook of AI Governance*. Oxford University Press.

Cabrol, N. A. (2016). *Alien mindscapes—a perspective on the Search for Extraterrestrial Intelligence*. Astrobiology, 16(9), 661-676.

Cassimatis, Nicholas L. (2006). *A cognitive substrate for achieving human-level intelligence*. AI magazine 27, no. 2: 45-45.

Coello, C. A. C. (2022). *Constraint-handling techniques used with evolutionary algorithms*. In Proceedings of the genetic and evolutionary computation conference companion (pp. 1310-1333).

Dagnall, N., Drinkwater, K., & Parker, A. (2011). *Alien visitation, extra-terrestrial life, and paranormal beliefs*. Journal of Scientific Exploration, 25(4), 699.

Floridi, Luciano (2019). *Establishing the rules for building trustworthy AI*. Nature Machine Intelligence 1:261-262.

Fourneret, Éric (2020). *The Hybridization of the Human with Brain Implants: The Neuralink Project*. Cambridge Quarterly of Healthcare Ethics 29 (4):668-672.

He, X., Zhao, K., & Chu, X. (2021). AutoML: A survey of the state-of-the-art. *Knowledge-based systems, 212*, 106622.

Hirvensalo, M. (2003). *Quantum computing*. Springer Science & Business Media.

Lindia, Matthew S. (2022). *Gadamer in a Wired Brain: Philosophical Hermeneutics and Neuralink*. Philosophy and Technology 35 (2):1-17.

Jabr, F., & Rothschild, A. (2012). How brainless slime molds redefine intelligence. *Scientific American, 7*, 1.

Kaplan, F. (2008). *Neurorobotics: an experimental science of embodiment*. Frontiers in neuroscience, 2, 1046.

Kar, M. K., Kumar, S., Singh, A. K., & Panigrahi, S. (2023). *Reactive power management by using a modified differential evolution algorithm*. Optimal Control Applications and Methods, 44(2), 967-986.

Konar, Amit. (2018). *Artificial intelligence and soft computing: behavioral and cognitive modeling of the human brain*. CRC press.

Marcus, Gary & Davis, Ernest (2019). *Rebooting Ai: Building Artificial Intelligence We Can Trust*. Vintage.

Michael, Haenlein & Kaplan, Andreas (2019). *A brief history of artificial intelligence: On the past, present, and future of artificial intelligence*. California management review 61.4: 5-14.

Miculan, M., & Peressotti, M. (2014). *GSOS for non-deterministic processes with quantitative aspects*. arXiv preprint arXiv:1406.2066.

Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Penguin UK.

Moustris, G. P., Hiridis, S. C., Deliparaschos, K. M., & Konstantinidis, K. M. (2011). Evolution of autonomous and semi-autonomous robotic surgical systems: a review of the literature. *The international journal of medical robotics and computer assisted surgery, 7*(4), 375-392.

Mullen, K. M. (2011). *Human-technology Integration*. In Unifying Themes in Complex Systems: Proceedings of the Fifth International Conference on Complex Systems (pp. 257-264). Berlin, Heidelberg: Springer Berlin Heidelberg.

M. Schlosser. (2019) Agency. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2019 edition.

Nick, B. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

O'Reilly, D., McGrath, J., & Martin-Loeches, I. (2023). Optimizing artificial intelligence in sepsis management: Opportunities in the present and looking closely to the future. *Journal of Intensive Medicine*.

Paulus, J. W., Nantes-Sobrinho, D., & Pérez, J. A. (2021). *Non-deterministic functions as non-deterministic processes*. In 6th International Conference on Formal Structures for Computation and Deduction (FSCD 2021). Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Sayed, G. I., & Hassanien, A. E. (2022). Explainable AI and Slime Mould Algorithm for Classification of Pistachio Species. *In Artificial Intelligence: A Real Opportunity in*

the Food Industry (pp. 29-43). Cham: Springer International Publishing.

Schürholt, K., Knyazev, B., Giró-i-Nieto, X., & Borth, D. (2022). Hyper-representations as generative models: Sampling unseen neural network weights. *Advances in Neural Information Processing Systems, 35*, 27906-27920.

Serag, A., Ion-Margineanu, A., Qureshi, H., McMillan, R., Saint Martin, M. J., Diamond, J., ... & Hamilton, P. (2019). Translational AI and deep learning in diagnostic pathology. *Frontiers in medicine, 6*, 185.

Shanker, Stuart (1995). *Turing and the origins of AI*. Philosophia Mathematica 3 (1):52-85.

Song, Y., Cai, X., Zhou, X., Zhang, B., Chen, H., Li, Y., ... & Deng, W. (2023). *Dynamic hybrid mechanism-based differential evolution algorithm and its application*. Expert Systems with Applications, 213, 118834.

Tegmark, M. (2018). *Life 3.0: Being human in the age of artificial intelligence*. Vintage.

Topal, A. O., Chitic, R., & Leprévost, F. (2023). *One evolutionary algorithm deceives humans and ten convolutional neural networks trained on ImageNet at image recognition.* Applied Soft Computing, 143, 110397.

van Lingen, Marlies N. ; Giesbertz, Noor A. A. ; van Tintelen, J. Peter & Jongsma, Karin R. (2023). *Why We Should Understand Conversational AI as a Tool*. American Journal of Bioethics 23 (5):22-24.

van Rijmenam, M., & Logue, D. (2021). *Revising the 'science of the organisation': Theorising AI agency and actorhood*. Innovation, 23(1), 127-144.

Williams, C. P. (2010). *Explorations in quantum computing*. Springer Science & Business Media.

Williams, M. D., & Lawson, P. C. (2020). *Analyzing the Challenges and Obstacles to Developing and Fielding Autonomous and Semi-Autonomous Systems* (Doctoral dissertation, Monterey, CA; Naval Postgraduate School).

Vesnic-Alujevic, L., Nascimento, S., & Polvora, A. (2020). *Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks*. Telecommunications Policy, 44(6), 101961.

Yu, X. and Gen, M., (2010). *Introduction to evolutionary algorithms*. Springer Science & Business Media.

Zhou, Aimin, Bo-Yang Qu, Hui Li, Shi-Zheng Zhao, Ponnuthurai Nagaratnam Suganthan, and Qingfu Zhang. (2011). *Multiobjective evolutionary algorithms: A survey of the state of the art*. Swarm and evolutionary computation 1, no. 1: 32-49.