Understanding Artificial Intelligence: A Multidisciplinary Analysis of AI's Distinct Cognitive Architecture

Giuseppe RIVA[a,b,d,1], Elena SAJNO[a,c], Stefano DE GASPARI[a,c], Chiara PUPILLO[a,c], Maria SANSONI[a,d], Giuseppe PASSALACQUA[a,c], Federico LONGONI [a,c]
Brenda K. WIEDERHOLD[e]

[a] *Humane Technology Lab, Catholic University of Sacred Heart, Milan, Italy*
[b] *Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, Milan, Italy*
[c] *Department of Computer Science, University of Pisa, Largo B. Pontecorvo, 3 56127 Pisa*
[d] *Department of Psychology, Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20100 Milan, Italy*
[e] *Virtual Reality Medical Center, Scripps Memorial Hospital, La Jolla, California, USA*
ORCiD ID: Giuseppe Riva https://orcid.org/0000-0003-3657-106X

Abstract. As Artificial Intelligence (AI) systems become increasingly sophisticated and pervasive, understanding their true nature and limitations requires moving beyond single-discipline perspectives. This paper argues that effectively developing and deploying AI systems demands an integrated framework combining insights from computer science, cognitive psychology, linguistics, and cyberpsychology. We examine three complementary theoretical frameworks that together reveal AI's fundamental nature: Psychomatics demonstrates how AI systems create sophisticated syntax-semantic mappings but lack true experiential understanding; System 0 thinking positions AI as a distinct cognitive layer that preprocesses information while remaining dependent on human interpretation; and Cyberpsychology illuminates how humans form mental models and relationships with AI systems. These insights converge on a central argument: effective AI development requires embracing rather than obscuring the fundamental differences between artificial and human intelligence. Instead of pursuing AI systems that attempt to replicate human cognition, we should focus on developing systems that transparently complement human capabilities while acknowledging their inherent limitations in meaning-making and understanding.

Keywords. Artificial intelligence, cyberpsychology, psychomatics, system 0

1. Introduction

The rapid evolution of artificial intelligence (AI) represents a complex intersection of multiple disciplines, each offering unique perspectives on how these systems process information and interact with human cognition. While AI, particularly Large Language Models (LLMs), demonstrates remarkable capabilities in language acquisition and concept learning, understanding its true nature requires a multidisciplinary lens [1] that spans computer science, cognitive psychology, linguistics, cyberpsychology, and ethics. Two emerging theoretical frameworks help illuminate this complex landscape. The "Psychomatics" framework, introduced by Riva and colleagues [2], bridges cognitive science, linguistics, and computer science to examine how LLMs acquire, learn, remember, and utilize information. This approach reveals how LLMs create extensive syntax-semantics maps to manipulate linguistic patterns while following Grice's Cooperative principle to generate relevant responses. However, unlike human cognition

[1] Corresponding Author: Giuseppe Riva; Humane Technology Lab., Catholic University of Sacred Heart, Milan, Italy; Email: giuseppe.riva@unicatt.it

Complementing this view, Chiriatti and colleagues [3] propose conceptualizing human-AI interaction as "System 0" - a distinct cognitive layer that operates alongside Kahneman's system 1 (fast, intuitive thinking) and system 2 (slow, analytical thinking). This framework positions AI systems as an artificial preprocessor that shapes how information flows into traditional human cognitive processes, while acknowledging AI's dependence on human interpretation for meaningful outputs.

Cyberpsychology plays a particularly crucial role in this interdisciplinary analysis by examining how individuals perceive and interact with AI systems. Research has shown that humans often anthropomorphize conversational AI, forming bonds that extend beyond mere tool usage. Understanding these psychological dynamics is essential for designing AI systems that are both user-friendly and ethically sound.

By adopting this multidisciplinary approach, we can foster a holistic understanding of AI that encompasses both its technical capabilities and its psychological, social, and ethical implications. Such a comprehensive perspective is vital for developing AI systems that are not only technologically advanced but also aligned with human values and societal needs.

2. Psychomatics: Understanding Digital Brains

Understanding how artificial intelligence processes information and generates responses is crucial for developing effective human-AI interactions. While LLMs demonstrate impressive capabilities in tasks ranging from language comprehension to problem-solving, their underlying mechanisms differ fundamentally from human cognition. The Psychomatics framework offers a systematic approach to understanding these differences by examining how LLMs acquire, process, and utilize information through the lens of cognitive science, linguistics, and computer science [1].

At its core, Psychomatics reveals that LLMs operate through extensive syntax-semantics mapping rather than genuine understanding. Unlike humans, who derive meaning from direct experience, emotional connections, and embodied cognition, LLMs create sophisticated statistical patterns that link linguistic structures to probable meanings. For example, when processing the sentence "The cat is on the mat," an LLM doesn't understand the physical reality of a cat or mat, but rather maps the statistical relationships between these terms and their contextual usage patterns. This mapping allows LLMs to generate coherent responses by following Grice's Cooperative principle [4] - providing relevant, informative, and appropriately structured communication - without truly comprehending the underlying concepts.

The framework identifies several key mechanisms that enable LLM functionality. Through self-attention mechanisms [5], LLMs capture dependencies between words within sequences, similar to how humans process syntagmatic relations in language. Cross-attention mechanisms allow LLMs to map relationships between different sequences, mimicking how humans form associative connections between concepts. For instance, in processing a medical query, an LLM might link symptoms to diagnoses not through medical understanding, but through statistical patterns learned from training data. However, this reliance on statistical patterns rather than embodied understanding leads to important limitations. LLMs can generate factually incorrect information (hallucinations) when their training data lacks sufficient context or when they encounter novel situations requiring real-world understanding. They struggle with implicit meanings, sarcasm, and contextual nuances that humans naturally grasp through their social and emotional experiences. Additionally, LLMs operate in an asocial environment, lacking the dynamic learning capabilities that humans develop through continuous social interaction and personal experience [6] .

The impact of LLMs on human-AI interaction is profound and multifaceted. While these systems can enhance human cognitive capabilities by processing vast amounts of information and generating sophisticated responses, they also risk creating a false sense of understanding [7]. Users might anthropomorphize LLMs and attribute deeper comprehension than actually exists, potentially leading to over-reliance or misuse of these systems. This highlights the importance of developing frameworks for responsible AI deployment that acknowledge both the capabilities and limitations of LLMs.

Looking ahead, the Psychomatics framework suggests that effective human-AI interaction requires understanding these fundamental differences between human and machine cognition. Rather than treating LLMs as human-like intelligence, we should view them as powerful but limited tools that complement human cognitive capabilities. This understanding can guide the development of AI systems that better serve human needs while maintaining appropriate boundaries and expectations. The future of human-AI interaction lies not in creating perfect simulations of human cognition, but in developing systems that transparently and effectively augment human capabilities while acknowledging their inherent limitations.

3. System 0: When Artificial Intelligence Augments Human Cognition

As artificial intelligence becomes increasingly integrated into our daily cognitive processes, understanding its relationship with human thought patterns has emerged as a critical frontier in cognitive science. While traditional frameworks like Kahneman's System 1 (fast, intuitive thinking) and System 2 (slow, analytical thinking) help us understand human cognition [8], they don't fully capture how AI systems are fundamentally reshaping our cognitive landscape. The concept of System 0 provides a theoretical framework for understanding this new cognitive reality [3], positioning AI as a foundational layer that preprocesses information before it enters human cognitive systems.

System 0 represents the outsourcing of cognitive tasks to AI systems that can process vast amounts of data and perform complex computations beyond human capabilities. Unlike Systems 1 and 2, which operate within individual minds, System 0 functions as an artificial, non-biological layer of distributed intelligence that interacts with and augments both intuitive and analytical thinking processes. For example, when using an AI writing assistant, the system doesn't merely extend our writing capabilities; it actively shapes how we approach the writing process by suggesting structures, identifying patterns, and proposing improvements before we engage in traditional cognitive processing. Similarly, when using AI-powered search engines, the information we receive is preprocessed and curated, fundamentally altering how we access and interpret knowledge.

The framework identifies several key characteristics that distinguish System 0 from traditional cognitive systems. It exhibits high information flow, with continuous two-way interaction between humans and AI. It demonstrates increasing reliability as AI systems become more sophisticated, leading to greater trust and procedural transparency in human-AI interactions. Perhaps most significantly, System 0 shows strong individualization capabilities, providing personalized cognitive support that adapts to individual users' needs and preferences.

However, System 0 differs crucially from Systems 1 and 2 in its lack of inherent meaning-making capabilities. While it can process and manipulate data with remarkable efficiency, it relies entirely on human interpretation to generate meaningful insights. This limitation creates a complex dynamic where human cognition must actively engage with AI-processed information to create genuine understanding.

The impact of System 0 on human decision-making is profound and potentially double-edged. On one hand, it enhances our cognitive capabilities by providing sophisticated preprocessing of information, enabling us to handle more complex problems and access broader knowledge bases. On the other hand, it raises important concerns about cognitive autonomy and critical thinking. As we increasingly rely on AI-mediated information and decision-making processes, we risk adopting epistemological norms aligned with computational logic rather than human reasoning.

Looking forward, the integration of System 0 into human cognition presents both opportunities and challenges. The potential benefits include enhanced problem-solving capabilities, more informed decision-making, and expanded cognitive reach. However, these must be balanced against the risks of decreased autonomy and potential erosion of critical thinking skills. The key to successful integration lies in developing frameworks that maintain human agency while leveraging AI's capabilities effectively.

Understanding System 0 is essential for developing AI systems that enhance rather than diminish human cognitive capabilities. This requires careful attention to how AI systems preprocess information, how humans interact with AI-generated insights, and how to maintain critical thinking skills in an AI-augmented cognitive environment. As we continue to develop and deploy AI systems, the concept of System 0 provides a valuable framework for ensuring these technologies serve as effective cognitive partners while preserving human autonomy and decision-making capabilities.

4. Cyberpsychology: The Psychology of Human-AI Relationships

As artificial intelligence becomes increasingly integrated into our daily lives, understanding how humans perceive and interact with AI systems has emerged as a critical frontier in technology research [9]. Cyberpsychology, the study of human interaction with technology, offers unique insights into the psychological dynamics that shape these human-AI relationships [10]. This field has become particularly relevant as AI systems evolve from simple tools to complex interactive agents that can engage in sophisticated dialogue, make decisions, and adapt to user behavior.

The psychological dimensions of human-AI interaction operate on multiple levels [11]. At the cognitive level, humans tend to anthropomorphize AI systems, attributing human-like characteristics and mental states to them even when knowing they are artificial. For example, research has shown that users often form emotional bonds with conversational AI, treating them as social entities rather than mere tools [12]. This phenomenon extends beyond simple projections - humans develop mental models of AI capabilities that significantly influence their trust, expectations, and willingness to rely on AI systems [13]. Studies demonstrate how these mental models can be either overly optimistic, leading to excessive trust in AI capabilities, or overly skeptical, resulting in resistance to beneficial AI applications [14]. The Psychomatics framework reveals how these perceptions are complicated by AI's inherent capability uncertainty and output complexity. Users must navigate systems whose abilities may evolve over time and whose outputs can range from simple to highly adaptive and complex. This uncertainty shapes how users build trust and develop usage patterns with AI systems.

The social-psychological aspects of human-AI interaction are equally important [15]. As highlighted by System 0 thinking, AI has become a foundational layer that preprocesses information before it enters human cognitive systems. This creates a complex dynamic where human decision-making and social behavior are increasingly mediated by AI systems, yet the psychological implications of this mediation are not fully understood. Users may experience cognitive offloading, where they become dependent on AI for tasks they previously handled independently, or develop new cognitive strategies that integrate AI capabilities into their problem-solving approaches. Cyberpsychology's impact on understanding AI interactions extends beyond theoretical insights to practical applications. By revealing how users develop mental models of AI systems, design teams can create more intuitive interfaces that properly calibrate user expectations [16]. Understanding the psychological mechanisms behind trust formation helps in developing AI systems that are not only technically capable but also psychologically accessible. Moreover, cyberpsychological research illuminates how different user groups may interact with AI differently based on their technological self-efficacy, cultural background, and previous experiences with technology [17].

Looking forward, cyberpsychology will play an increasingly crucial role in shaping the future of human-AI interaction. As AI systems become more sophisticated and pervasive, understanding the psychological dimensions of these interactions will be essential for developing technologies that enhance rather than diminish human capabilities. This understanding can help address critical challenges in AI deployment, from preventing over-reliance on AI systems to ensuring that AI augments human decision-making in psychologically healthy ways. The field's insights will be particularly valuable in sensitive domains like healthcare, education, and personal assistance, where the psychological dynamics of human-AI interaction directly impact system effectiveness and user wellbeing.

5. Conclusions

The multidisciplinary analysis presented in this paper leads to several critical insights about the nature of artificial intelligence and its relationship to human cognition. First, the Psychomatics framework reveals that while AI systems can create sophisticated statistical patterns linking linguistic structures to probable meanings, they fundamentally lack the experiential and embodied understanding that underlies human cognition. This isn't merely a temporary limitation but reflects a fundamental difference in how AI systems process and derive meaning from information.

Second, the System 0 perspective demonstrates that AI systems are best understood not as autonomous thinking entities but as a distinct cognitive layer that preprocesses information for human interpretation. This positioning helps resolve the apparent paradox of AI systems that can perform complex cognitive tasks while lacking true understanding - they function as sophisticated information processors that augment rather than replicate human cognitive capabilities.

Third, cyberpsychology research shows how humans naturally form mental models and relationships with AI systems that often overestimate their capabilities for genuine understanding. This tendency toward anthropomorphizing AI needs to be carefully managed to ensure appropriate trust and reliance on AI systems.

These insights converge on a central argument: effective AI development requires embracing rather than obscuring the fundamental differences between artificial and human intelligence. Instead of pursuing AI systems that attempt to replicate human cognition, we should focus on developing systems that transparently complement human capabilities while acknowledging their inherent limitations in meaning-making and understanding. This approach has several practical implications:

- AI system design should emphasize transparency about operational limitations rather than attempting to create the illusion of human-like understanding;
- Development frameworks should focus on creating effective human-AI partnerships that leverage the complementary strengths of both;
- Implementation strategies should include explicit consideration of how humans form psychological relationships with AI systems.

Looking ahead, this integrated understanding provides a foundation for developing AI systems that enhance human cognitive capabilities while maintaining appropriate boundaries and expectations. Success in this endeavor requires continued cross-disciplinary dialogue and research that brings together technical capabilities, cognitive science insights, and psychological understanding of human-AI interaction.

6. Acknowledgments

This paper was funded by the European Union - Next Generation EU - Italian PNRR under the FAIR Project "Co-XAI - Cognitive Decision Intelligence Framework for Explainable AI Systems". The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

[1] Dignum V. AI is multidisciplinary. AI Matters. 2020;5:18–21. doi: https://doi.org/10.1145/3416666.3421426.

[2] Riva G, Mantovani F, Wiederhold BK, Marchetti A, Gaggioli A. Psychomatics—A multidisciplinary framework for understanding artificial minds. Cyberpsychol Behav Soc Netw. 2024. doi: https://doi.org/10.1089/cyber.2024.0003.

[3] Chiriatti M, Ganapini M, Panai E, Ubiali M, Riva G. The case for human–AI interaction as system 0 thinking. Nat Hum Behav. 2024;8:1829–1830. doi: https://doi.org/10.1038/s41562-024-01872-7.

[4] Grice HP. Logic and conversation. In: Cole P, Morgan JL, editors. *Syntax and Semantics: Vol 3. Speech Acts.* New York: Academic Press; 1975. p. 41–58.

[5] Vaswani A, Shazeer NM, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is all you need. In: *Advances in Neural Information Processing Systems*; 2017. p. 5998–6008.

[6] Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. J Med Internet Res. 2019;21:e13216. doi: https://doi.org/10.2196/13216.

[7] Messeri L, Crockett MJ. Artificial intelligence and illusions of understanding in scientific research. Nature. 2024;627:49–58. doi: https://doi.org/10.1038/s41586-024-07372-3.

[8] Kahneman D. Maps of bounded rationality: a perspective on intuitive judgment and choice. In: Frängsmyr T, editor. *The Nobel Prizes 2002.* Stockholm: Nobel Foundation; 2002. p. 449–489.

[9] Xu W, Dainoff MJ, Ge L, Gao Z. Transitioning to human interaction with AI systems: new challenges and opportunities for HCI professionals to enable human-centered AI. Int J Hum Comput Interact. 2023;39(5):494–518. doi: https://doi.org/10.1080/10447318.2022.2094212.

[10] Fortuna P. Positive cyberpsychology as a field of study of the well-being of people interacting with and via technology. Front Psychol. 2023;14:1237890. doi: https://doi.org/10.3389/fpsyg.2023.1237890.

[11] Hancock PA, Billings DR, Schaefer KE, Chen JYC, de Visser EJ, Parasuraman R. A meta-analysis of factors affecting trust in human-robot interaction. Hum Factors. 2011;53:517–527. doi: https://doi.org/10.1177/0018720811417254.

[12] Pentina I, Hancock T, Xie T. Exploring relationship development with social chatbots: a mixed-method study of Replika. Comput Hum Behav. 2023;140:107600. doi: https://doi.org/10.1016/j.chb.2022.107600.

[13] Kouros T, Papa V. Digital mirrors: AI companions and the self. Societies. 2024;14(2):200. doi: https://doi.org/10.3390/soc14020024.

[14] Pan W, Liu D, Meng J, Liu H. Human–AI communication in initial encounters: how AI agency affects trust, liking, and chat quality evaluation. New Media Soc. 2024;14614448241259149. doi: https://doi.org/10.1177/14614448241259149.

[15] Bingley WJ, Haslam SA, Steffens NK, Gillespie N, Worthy P, Curtis C, et al. Enlarging the model of the human at the heart of human-centered AI: a social self-determination model of AI system impact. New Ideas Psychol. 2023;70:101025. doi: https://doi.org/10.1016/j.newideapsych.2023.101025.

[16] Pataranutaporn P, Liu R, Finn E, Maes P. Influencing human–AI interaction by priming beliefs about AI can increase perceived trustworthiness, empathy and effectiveness. Nat Mach Intell. 2023;5:1076–1086. doi: https://doi.org/10.1038/s42256-023-00750-0.

[17] Andrews RW, Lilly JM, Srivastava D, Feigh KM. The role of shared mental models in human-AI teams: a theoretical review. Theor Issues Ergon Sci. 2023;24(2):129–175. doi: https://doi.org/10.1080/1463922X.2023.2165766.