*Global Healthcare Transformation in the Era of Artificial Intelligence and Informatics*
*J. Mantas et al. (Eds.)*

*doi:10.3233/SHTI250690*

# Relevance of Grounding AI for Health Care

Murat SARIYAR[a,1]
[a]*Bern University of Applied Sciences, Switzerland*
ORCiD ID: Murat Sariyar https://orcid.org/0000-0003-3432-2860

**Abstract.** As large language models (LLMs) like GPT-4 are increasingly deployed in clinical and administrative healthcare settings, questions about their conceptual grounding take on renewed urgency. While concerns about the lack of sensorimotor experience in symbolic AI systems have been long discussed in cognitive science and philosophy of mind, their practical implications in medicine remain underexplored. This paper revisits the grounding problem through the lens of contemporary healthcare applications, arguing that the unique demands of medical reasoning – interpretive nuance, ethical sensitivity, and contextual depth—amplify the limitations of ungrounded AI. By reframing classic debates, such as Searle's Chinese Room and the Symbol Grounding Problem, within real-world clinical contexts, we highlight specific risks that emerge when LLMs are treated as epistemic agents rather than tools.

**Keywords.** Artificial intelligence (AI), artificial general intelligence (AGI), Large Language Model (LLMs), symbol grounding.

## 1. Introduction

The role of AI in healthcare is expanding rapidly, encompassing applications ranging from diagnostic support to the optimization of administrative processes [1]. However, most current AI systems, including large language models (LLMs) such as GPT-4, operate solely through symbolic manipulation, lacking any grounding in real-world experience [2]. This lack of grounding – defined as the absence of sensorimotor interaction and subjective understanding – raises important concerns about the reliability and limitations of AI in high-stakes domains like healthcare [3]. Whether symbolic grounding is essential for effective AI performance in such contexts remains an open question. And if grounding is indeed necessary, identifying viable approaches to achieve it poses a significant challenge, especially given that even in humans, the nature of understanding is not yet fully understood.

"Grounding" is a philosophical concept asserting that the symbols must be rooted in direct experience to be genuinely understood [4]. Without a non-symbolic experiential layer, it is impossible to transcend the closed loop of symbolic reference. It is not enough for someone else to have the experience and convey its symbolic representations to me. Mere symbolic manipulation or rote memorization does not constitute true understanding[5]. Thus, clarifying the relationship between grounding and understanding is essential, especially related to artificial general intelligence (AGI).

[1] Corresponding Author: Murat Sariyar, Bern University of Applied Sciences, Quellgasse 21, CH2502 Biel/Bienne, Switzerland; E-mail: murat.sariyar@bfh.ch.

This paper explores the conceptual foundations of grounding in AI and its implications for healthcare applications. It draws on key theoretical frameworks, including Searle's Chinese Room Argument (CRA [6]), the Symbol Grounding Problem (SGP [7]), and the distinctions between T2 and T3 AI capabilities [8]. By analyzing these theories, the paper aims to elucidate the limitations of ungrounded systems and evaluate their relevance in healthcare contexts. For instance, does a chatbot need to "understand" that a patient is suicidal, or is it sufficient for the system to recognize patterns that lead to the assessment of "suicidal", along with its associated consequences?

## 2. Methods

### *2.1. Language, Meaning, and Understanding*

Language processing inherently involves meaning and understanding, but these aspects manifest in two distinct dimensions. Symbolic Manipulation: LLMs like GPT-4 excel at generating coherent language by processing and manipulating symbols. This is what Searle terms the "easy" problem of understanding: the ability to mimic meaningful interaction through syntactic operations [9]. Phenomenological Understanding: The "hard" problem involves subjective, felt experiences – what it "feels like" to understand a concept. This aspect is entirely absent in current AI systems. In healthcare, effective communication often requires more than syntactic fluency; it demands a deep understanding of context, patient needs, and the implications of medical information. For example, correctly interpreting the phrase "severe allergic reaction to medication" requires not only linguistic competence but an understanding of its real-world implications for treatment decisions.

### *2.2. Symbol Grounding and Sensorimotor Interaction*

The Symbol Grounding Problem (SGP) highlights a core limitation of many current AI systems: the inability of symbols to acquire intrinsic meaning without connection to real-world, sensorimotor experiences. In healthcare, for example, understanding the concept of fever involves more than recognizing the word's linguistic patterns – it requires grounding in sensory and conceptual experiences like elevated body temperature, patient discomfort, and associated medical conditions. Most AI systems today rely on indirect grounding. They infer meaning through statistical associations in text corpora rather than through interaction with the physical world. These are what Harnad terms T2 systems, which pass the Turing Test by producing linguistically coherent outputs that are indistinguishable from those of a human conversational partner [9]. This test, often described as a "pen-pal test," evaluates only disembodied verbal performance and omits any requirement for sensorimotor engagement.

In contrast, the Total Turing Test (T3) introduces a more stringent criterion. A T3 system must not only demonstrate linguistic fluency but also the ability to perceive, manipulate, and interact with real-world objects and situations. This form of sensorimotor grounding ties symbols to their referents through lived experience rather than solely through linguistic co-occurrence. Such grounding is especially critical in domains like medicine, where decision-making depends on the interpretation of physical signs, environmental cues, and embodied interactions with patients [10]. Yet even passing the T3 test does entail consciousness or genuine subjective understanding. Pushing further, the T4 test – which posits complete neurobiological equivalence – still fails to bridge the explanatory gap between externally observable behavior and the internal reality of conscious experience. These distinctions (see also Figure 1) underscore a fundamental limitation in cognitive science: no matter how advanced an AI system becomes; we cannot fully verify subjective experience through objective tests alone. In high-stakes fields like healthcare, this means that trust in AI systems ultimately rests on their observable reliability and contextual appropriateness, not on any claim to genuine understanding. Symbol grounding, then, is not merely a philosophical issue but a practical concern for how and where AI can safely be deployed.

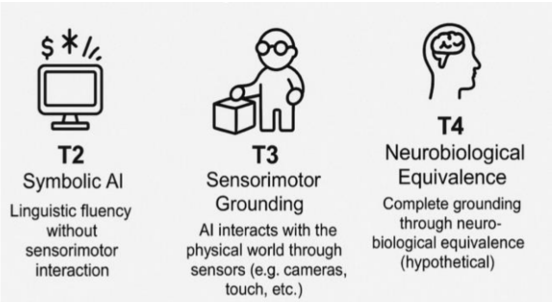


**Figure 1.** Illustration of AI Grounding Levels: T2 (Symbolic AI), T3 (Sensorimotor Grounding), and T4 (Neurobiological Equivalence).

## 3. Results

In applied medical AI, the central concern is not whether artificial systems possess human-like understanding, but whether they can simulate it well enough to support safe and effective clinical decision-making. Hence, the philosophical resolution of the SGP may not be essential in theory, but its practical implications become still critical when AI is deployed in real-world contexts. Here, "understanding" refers to the operational ability to recognize facts, situate them in context, and infer clinically relevant implications. Consider a system interpreting the phrase "The patient is febrile." An ungrounded model may statistically associate "febrile" with "fever" and generate plausible differential diagnoses, but it cannot infer urgency based on patient age, immunosuppression, or symptom clusters. This limits their ability to form integrated, situational models essential for nuanced clinical reasoning – capabilities rooted in human cognition, which continuously synthesizes sensory data with memory, context, and intent.

Even multimodal AI models, which combine text with images or other modalities, struggle to link perceptual data to medical semantics in contextually robust ways. A model might label an image of a skin rash as "urticaria," but fail to factor in progression, anatomical distribution, or patient-reported symptoms – critical for diagnosis. Lacking sensorimotor grounding, such systems cannot disambiguate observations or verify outputs against real-world causality, introducing epistemic opacity. This opacity becomes dangerous in clinical domains, where a plausible but unverified output can lead to harmful decisions. While techniques like reinforcement learning or sensor integration improve contextualization, they still rely on predefined features and human-curated signals. The combinatorial complexity of clinical practice – where conditions present variably, and context often defines urgency – makes it infeasible to anticipate all relevant variables during training. Consequently, generalization beyond narrow domains remains fragile, with a growing risk of silent failure in untested scenarios.

One striking example is the use of AI chatbots in mental health. These systems can mimic therapeutic language and produce superficially empathetic responses but lack emotional and situational grounding essential for assessing psychological crises. An AI might identify "I feel hopeless" as a suicide risk indicator and return a templated reply, while missing sarcasm, history, or emotional nuance. Such misinterpretations may lead to ineffective or harmful recommendations. Unlike clinicians, who draw on experience, embodied empathy, and intuition, AI systems remain contextually brittle black boxes. Even T3 systems lack the phenomenological and interpretive capacities inherent in human cognition. Designing AI to complement human expertise – rather than imitate it – offers a more realistic and ethically sound path forward.

This perspective suggests domain-specific adaptations rather than the pursuit of general, autonomous intelligence. LLMs exemplify this approach, where despite their ungrounded nature, they can be optimized for medical tasks by aligning their outputs with domain-specific norms, such as diagnostic guidelines and clinical notes. However, like medical devices, the safe integration of AI in healthcare depends on rigorous validation, not assumed competence. Responsible use requires domain-specific constraints, performance guarantees, and continuous oversight to ensure reliability in real-world applications. To further address the limitations of T2 systems, improvements such as enhanced patient simulation environments and the integration of diverse data sources – like medical history and environmental factors – can better contextualize AI-driven decisions and improve their accuracy in complex clinical settings.

## 4. Discussion and Conclusions

A key challenge in medical AI development is that grounding is not just an academic issue but a practical necessity. Without grounding, AI struggles to interpret complex medical scenarios, where meaning is shaped by both textual cues and sensorimotor or environmental data. For example, interpreting "shortness of breath" requires understanding its physical signs, progression, and context. Integrating sensorimotor interactions with language processing is a promising approach for creating more robust and context-aware AI systems. However, AI's grounding diverges from human grounding, which emerges through lifelong learning, embodiment, and causal inference. Humans understand through exploration, feedback, and intrinsic motivation, learning to connect perception, action, and consequence. AI, in contrast, relies on static datasets and predefined goals, lacking mechanisms for open-ended learning or context retention. To approximate human-like reasoning, AI must develop interactive learning, spatiotemporal inference, and adaptive generalization – capacities that remain unrealized. Achieving true perceptual and conceptual grounding is a critical open question in AI research.

While fully grounded AI systems are not yet a reality, hybrid models like multimodal LLMs, which integrate text with visual or sensor-based data, show promise in bridging the grounding gap. However, their lack of lived experience limits their ability to understand context in dynamic environments like healthcare. To ensure safe deployment in high-stakes settings, ethical guardrails are crucial. Transparency alone is insufficient; AI must exhibit predictable, controllable behavior aligned with harm-reduction principles. This includes real-time escalation protocols, human-in-the-loop feedback, and embedded ethical constraints. Additionally, safety measures like automated handoffs to human clinicians in high-risk situations are essential to mitigate harm and ensure AI supports, rather than destabilizes, healthcare practices.

In parallel, the evolution of AI in healthcare calls for regulatory modularity – a flexible oversight architecture that can evolve with system capabilities [11]. Traditional validation pipelines, typically designed for static software or hardware, are poorly suited to adaptive, learning-based AI, which continues to evolve post-deployment. A modular regulatory framework would enable task-specific certification checkpoints, ensuring AI systems remain aligned with safety and performance standards as they evolve. For instance, an AI trained to detect pulmonary nodules could be validated for that task, then re-evaluated before expanding to other domains like cardiovascular or neurologic abnormalities. This approach allows scalable growth while maintaining compliance and supports iterative improvement, acknowledging the dynamic nature of AI systems. However, broader trust in AI beyond modular frameworks requires solving the symbol grounding problem to ensure true contextual understanding.

In summary, the future of AI in healthcare hinges on a multi-faceted strategy that combines sensorimotor grounding, adaptive architecture, and ethical design. Progress requires not only technical advances but also new regulatory models and conceptual frameworks that emphasize collaboration, constraint, and contextual awareness. Rather than aiming for full autonomy, AI should be developed as a cooperative agent – an extensible layer of cognitive support that enhances the precision and reach of human expertise. With such an approach, AI can become a powerful, safe, and accountable tool in medicine – amplifying capabilities while minimizing risks of misalignment and harm.

## References


[1] Khalifa M, Albadawy M, Iqbal U. Advancing Clinical Decision Support: The Role of Artificial Intelligence Across Six Domains. Comput Methods Programs Biomed Update 2024; 5: 100142.

[2] Pavlick E. Symbols and Grounding in Large Language Models. Philos Trans R Soc Math Phys Eng Sci 2023; 381: 20220041.

[3] Mumuni A, Mumuni F. Large Language Models for Artificial General Intelligence (AGI): A survey of Foundational Principles and Approaches. Epub ahead of print 6 January 2025. DOI: 10.48550/arXiv.2501.03151.

[4] Li J, Mao H. The Difficulties in Symbol Grounding Problem and the Direction for Solving It. Philosophies 2022; 7: 108.

[5] Abbate F. Natural and Artificial Intelligence: A Comparative Analysis of Cognitive Aspects. Minds Mach 2023; 33: 791–815.

[6] Lengbeyer L. Dismantling the Chinese Room with Linguistic Tools: a Framework for Elucidating Concept-Application Disputes. AI Soc 2022; 37: 1625–1643.

[7] Nagoev Z, Nagoeva O, Anchokov M, et al. The Symbol Grounding Problem in the System of General Artificial Intelligence based on Multi-Agent Neurocognitive Architecture. Cogn Syst Res 2023; 79: 71–84.

[8] Harnad S. Minds, Machines and Turing. In: Moor JH (ed) The Turing Test: The Elusive Standard of Artificial Intelligence. Dordrecht: Springer Netherlands, 253–273.

[9] Harnad S. Language Writ Large: LLMs, ChatGPT, Grounding, Meaning and Understanding. Epub ahead of print 3 February 2024. DOI: 10.48550/arXiv.2402.02243.

[10] French RM. Peeking Behind the Screen: the Unsuspected Power of the Standard Turing Test. J Exp Theor Artif Intell 2000; 12: 331–340.

[11] Widder DG, Nafus D. Dislocated accountabilities in the “AI supply chain”: Modularity and Developers’ Notions of Responsibility. Big Data Soc 2023; 10: 20539517231177620.