

# Exploring the Uncanny: AI Hallucinations and the Intriguing Intersection of Technology and Perception

By: Thamsanqa Sibiyi  
Founder and CEO of Siza Health

---



**Fig.1 AI Hallucination Depiction, Bernard Marr, 2023**

In the realm of artificial intelligence (AI), we often marvel at its ability to perform complex tasks, make predictions, and even emulate human creativity. However, there's a fascinating and slightly eerie aspect of artificial intelligence that has been gaining attention in recent years. These digital phantoms blur the lines between reality and virtuality, offering a glimpse into the mysterious world where technology and perception converge.

## **The Enigma of AI Hallucinations**

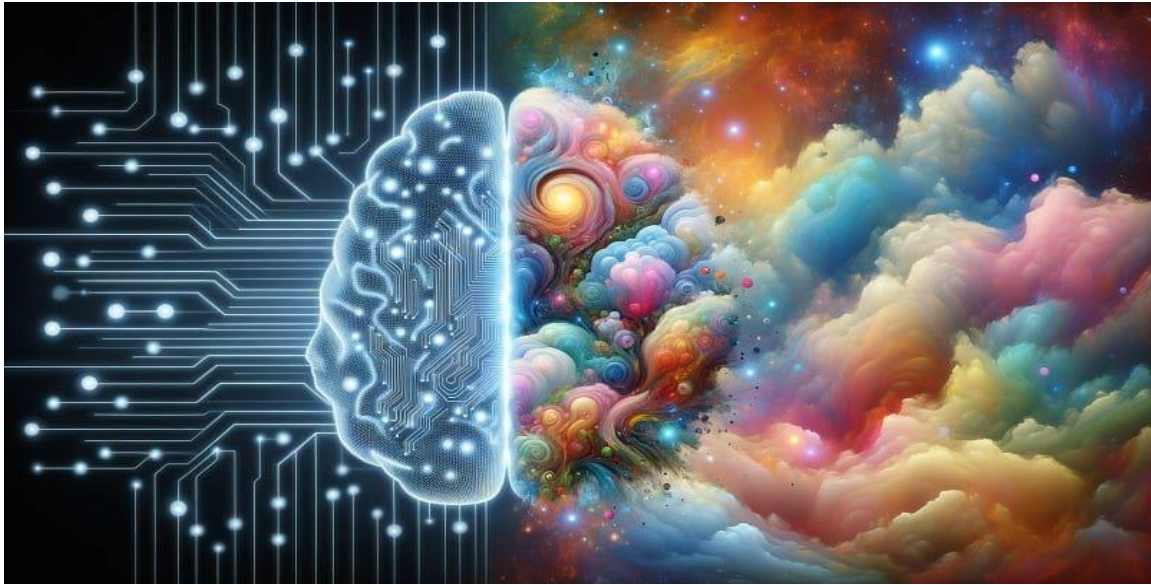


Fig.2 Enigma of Hallucinations, DALL-E 2023

Hallucinations in artificial intelligence, in the context of large language models occur when the models generate results / outputs that are not grounded in reality. They occur when algorithms generate unexpected and surreal outputs that seem to transcend their programming. These digital mirages can take various forms, from bizarre images and dreamlike landscapes to uncanny audio experiences. While these hallucinations are not intentional creations, they highlight the intricate nature of AI systems and the enigmatic ways in which they interpret and process information.



Fig.3 Google's DeepDream depiction of Vincent Van Gogh, Deep Dreamer Generator, 2016

One of the earliest instances of AI hallucinations emerged from Google's DeepDream, a neural network designed for image recognition in 2015 although the phenomenon started much earlier. DeepDream's algorithm, when turned inwards, started producing hallucinatory images filled with strange creatures and intricate patterns. What began as a simple experiment in image recognition soon transformed into an artistic exploration of the AI's imagination.

Hussam Alkaissi and Samy I. McFarlane, noted in their 19<sup>th</sup> February 2023 paper, **Artificial Hallucinations in ChatGPT: Implications in Scientific Writing** that, when they asked ChatGPT to provide a short paragraph on the mechanism of homocysteine-induced osteoporosis, “it touched on three main aspects, osteoblast inhibition, osteoclasts over activity, and, surprisingly, their mechanism on vitamin K-related carboxylation of osteocalcin”.

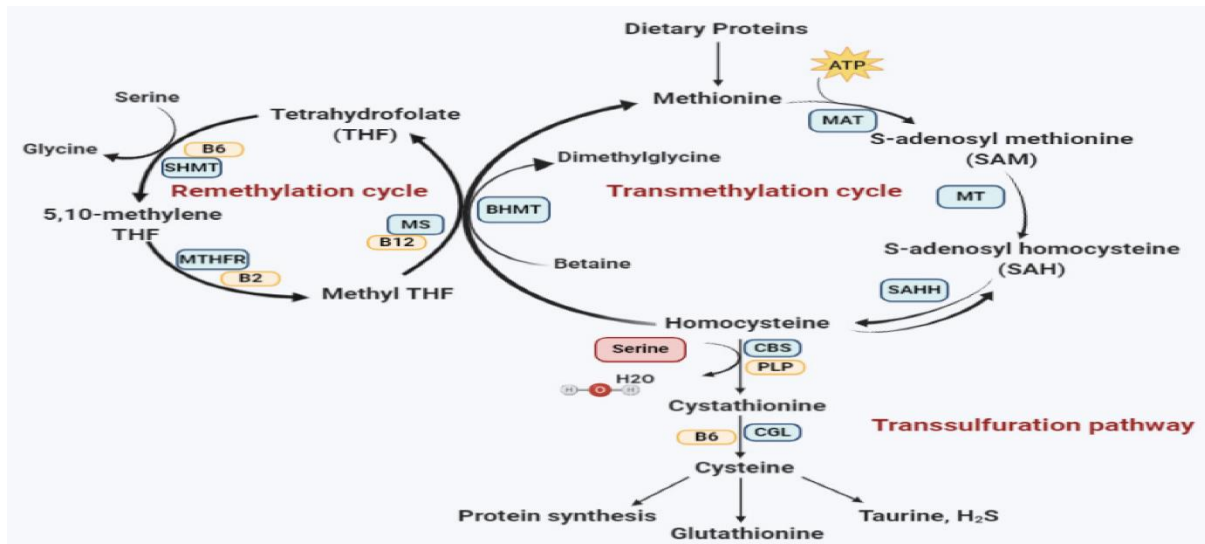


Fig.4 The Spectrum of Mutation of Homocystinuria, Genes (Basel) On-Line, 2020

Thorough review of the literature on bone metabolism and homocysteine shows that, the first two facts provided by ChatGPT were correct regarding osteoblast and osteoclast imbalance and the progression of osteoporosis.

“Similarly, when taken alone, the biochemistry of undercarboxylated osteocalcin and osteoporosis is the valid mechanism by which vitamin K deficiency is associated with osteoporosis. Homocysteine can reduce osteocalcin production but has nothing to do with post-translational carboxylation of osteocalcin glutamate residues”, writes Hussam Alkaissi and Samy I. McFarlane, 2023, 2

When references dating back over 20 years that had been provided by ChatGPT were checked,” none of the provided paper titles existed, and all provided PubMed IDs (PMIDs) were of different unrelated papers”. "Kallajoki M, et al.





Fig.5 Illustration of Google's WaveNet, agsandrew (opens in a new window)/ Shutterstock, 2016

Another example is that of Google's WaveNet, an AI system for generating realistic human-like speech, has been known to create hauntingly beautiful melodies and eerie soundscapes. When researchers experimented with the parameters of WaveNet, it resulted in unexpected audio hallucinations, showcasing the AI's capacity to explore realms beyond its intended functionality.



Fig.6 Ethical implications of AI hallucination, Artemis Diana (Getty Images)

This leads us to the Implications and Ethical Considerations that while the advancement in LLMs offer a captivating glimpse into the potential of artificial intelligence, they also raise important questions about ethics and control. As we delegate more decision-making to AI systems, the unintended consequences of these hallucinations underscore the need for responsible development and monitoring. These hallucinations produced by AI systems may inadvertently reflect the biases present in the training data. This phenomenon serves as a reminder of the importance of diverse and representative datasets to mitigate the risk of perpetuating societal biases through AI outputs.



Fig.7 The future impact of AI and the potential it holds for humanity, Getty Images

The surreal outputs of AI hallucinations challenge our conventional understanding of reality, and the potential AI holds for the future. As AI increasingly becomes a part of our daily lives, these digital **mirages** prompt us to consider how they might influence our perceptions and shape our cultural narratives.

AI hallucinations provide a captivating glimpse into the untamed creativity of artificial intelligence. Most leading lights in the artificial intelligence space, such as Steve Wozniak and Elon Musk have called for a slow down or even a total pause on all artificial intelligence research until such time that the risks of artificial hallucinations are understood and can be mitigated. Interestingly, while the CEO of Open AI acknowledges the existence of AI

Hallucinations, he offers a rather different view in that he sees these as “systems coming up with new ideas and being creative”.

Stuart J. Russel, a leading British computer scientist and an artificial intelligence researcher goes further by saying that AI systems spread misinformation and will cause untold destruction if placed in the wrong hands.

As we navigate this uncharted territory, it is crucial to approach these phenomena with curiosity and caution. By understanding the mechanisms behind AI hallucinations, we can harness the potential of technology while also addressing the ethical implications that arise in this fascinating intersection of technology and perception. As AI continues to evolve, the exploration of its hallucinatory capabilities opens new doors to artistic expression, philosophical contemplation, and a deeper understanding of the intricate dance between machines and the human mind. Artificial Intelligence Hallucinations challenges us to accept that we cannot leave the fate of humanity to machines. We are called to accept that at every step of the way, these machines require of us that we supervise to avoid disasters.