THE EPISTEMIC LIMITS OF AI IN VUCA CONTEXTS: A CRITIQUE OF SEMANTIC CLUSTERING, VECTORIZATION, AND TOKENIZATION

[1]**Harsh Patel**

[1]B.Tech. Aerospace Eng. (IIT Kharagpur)

ABSTRACT:

This paper explores the epistemological limitations of AI, particularly Large Language Models (LLMs) and Natural Language Processing (NLP) systems that rely on semantic clustering, vectorization, and probabilistic tokenization. While these techniques allow AI to generate linguistically coherent outputs, they fail to replicate human cognition, which is grounded in tacit knowledge, conceptual frameworks, and causal reasoning. Drawing from Polanyi's notion of tacit knowledge and Kant's categories of understanding, this study argues that AI's reliance on association-based processing inherently limits its ability to engage in counterfactual reasoning and dynamic decision-making.

These epistemic constraints have practical consequences, particularly in VUCA (Volatile, Uncertain, Complex, and Ambiguous) business environments. AI thrives in structured, data-intensive tasks but struggles with uncertainty, strategic foresight, and adaptability—domains where human expertise remains indispensable. Thus, this paper advocates for a human–AI hybrid approach, where AI enhances efficiency but does not replace human judgment in critical decision-making.

Keywords: Large Language Models (LLMs), Natural Language Processing (NLP), semantic clustering, vectorization, probabilistic tokenization, epistemic limitations, tacit knowledge, counterfactual reasoning, causal inference, AI in VUCA environments, volatility, uncertainty, complexity, ambiguity, transformer architectures, tokenization methods, stochastic parroting, automation risks, business intelligence, strategic foresight, AI-human hybrid models, ethical AI, crisis management, decision-making, governance, structured vs. unstructured tasks, knowledge representation, automation of cognitive tasks, AI epistemology, AI as a decision-support tool, human-AI collaboration, sustainability in AI integration

Methodology:

This research employs a **conceptual and theoretical analysis**, drawing from:

- **Epistemology & Cognitive Science**: Examining **Polanyi's and Kant's** frameworks to highlight AI's inability to capture tacit knowledge.
- **Mechanistic Deconstruction of LLMs**: Analyzing AI's reliance on **vectorization, token prediction, and semantic clustering** to illustrate its epistemic constraints.
- **Case-Based Evaluation**: Investigating **business failures of AI in VUCA contexts**, such as crisis management and strategic planning.
- **Implications for Sustainable AI Integration**: Recommending strategies for businesses to optimize AI's strengths while **preserving human adaptability and reasoning**.

Structural Epistemology of LLMs: Parsing, Vectorization, and Tokenization:

LLMs (Large Language Models) process linguistic data using multi-level deep neural networks, with recent models predominantly leveraging Transformer architectures that

ISSN No: 2349-059X
Peer Reviewed & Refereed Journal
National Research Journal

implement self-attention mechanisms. Unlike earlier Statistical and Neural Language Models, Transformers enable context-aware vectorization, improving semantic clustering and probabilistic token prediction. LLMs are trained on vast datasets of information across the internet and have their own tokenization methods as per their models. It could be a word, sub-word or character level tokenization. Essentially, all sentences and words in a language that the model has access to over the internet and other pre-training data, are broken into manageable chunks called tokens. These models do not "understand" language; they process numerical representations of tokens, mapping words into high-dimensional spaces based on statistical correlations rather than conceptual meaning.

To illustrate, consider an LLM trained on the following sentences:

1. "The cat sat on the mat."
2. "The dog lay on the floor."
3. "The cat chased the mouse."

Let’s also assume that after tokenization at the word level, the dataset transforms into a sequence: ["The", "cat", "sat", "on", "the", "mat", ".", "The", "dog", "lay", "on", "the", "floor", ".", "The", "cat", "chased", "the", "mouse", "."] Let’s focus on the three words: “cat”, “mat” and “chased”. The LLM embeds each word as a three-dimensional vector based purely on statistical correlations, without inherent meaning. For clarity, assume these dimensions roughly represent ‘Animalness,’ ‘Action Tendency,’ and ‘Location Tendency’—though the model itself does not interpret them this way. We denote these word embeddings as V vectors.

Word	V (Value) Vector Representation
"cat"	[0.8, 0.3, 0.6]
"mat"	[0.2, 0.4, 0.5]
"chased"	[0.5, 0.5, 0.8]

These numbers are not derived from logic or understanding, which is why “chased” may score higher than “mat” on the ‘Animalness’ metric or higher than “cat” on ‘Location tendency.’ The LLM does not comprehend these words in conceptual terms but merely assigns vectors based on statistical associations. For instance, the word “dog” might be vectorized as [0.6, 0.4, 0.9], not because the model understands its attributes, but because it frequently appears in similar contexts to “cat” rather than “mat” or “floor.”

Modern Transformer-based LLMs refine these vector representations using self-attention, feedforward networks, and backpropagation. Self-attention mechanisms determine how words relate to each other in a given context. Each word is mapped to a Query (Q) and a Key (K) vector, which dictate what information it seeks and how it responds. It does this via weight matrices which are improved through continuous learning. Let us assume that the learned weight matrices (W_q, W_k and W_v) are as follows:

W_q = [[0.5, 0.2, 0.1],
[0.3, 0.7, 0.2],
[0.6, 0.4, 0.8]]

W_k = [[0.4, 0.1, 0.3],
[0.8, 0.5, 0.6],
[0.2, 0.7, 0.9]]

W_v = [[0.9, 0.2, 0.5],
[0.3, 0.8, 0.6],
[0.5, 0.4, 0.7]]

For "cat", the Q, K and revised V Vectors can be obtained as follows:

$Q_i = W\_q_{ij}.V_j$

$K_i = W\_k_{ij}.V_j$

$V_i = W\_v_{ij}.V_j$

which would lend us Q(cat) = [0.52, 0.57, 1.08], K(cat) = [0.53, 1.15, 0.91] and V(cat) (revised) = [1.08, 0.84, 0.94]. Similarly, we get Q(mat) = [0.23, 0.44, 0.68], K(mat) = [0.27, 0.66, 0.77], V(mat) (revised) = [0.46, 0.68, 0.61] and Q(chased) = [0.43, 0.66, 0.9], K(chased) = [0.49, 1.13, 1.17], V(chased) (revised) = [0.95, 1.03, 1.01].

To enhance semantic clustering, the Transformer computes dot product attention scores between words, determining their contextual relevance. Specifically, it evaluates how the word "cat" attends to "mat" and "chased" by taking the dot product of the Query Vector Q(cat) with the Key Vectors K(mat) and K(chased).

Attention(cat → mat) = Q(cat).K(mat) = [0.52, 0.57, 1.08] . [0.27, 0.66, 0.77] = 1.3482

Attention(cat → chased) = Q(cat).K(chased) = [0.52, 0.57, 1.08] . [0.49, 1.13, 1.17] = 2.1625

Since 2.1625 > 1.3482, the model assigns stronger attention to "chased" than to "mat." The attention weights are then normalized using softmax:

Weight(cat → mat) = exp(1.3482) / [exp(1.3482) + exp(2.1625)] ~ 0.31

Weight(cat → chased) = exp(2.1625) / [exp(1.3482) + exp(2.1625)] ~ 0.69

Let's capture these Attention weights from "cat" to the other words by W_a(cat). Also, let's denote the two Value Vectors V(mat) and V(chased) into a higher order Value Vector V(mat, chased). Thus, W_a(cat) = [Weight(cat → mat), Weight(cat → chased)] and V(mat, chased) = [V(mat), V(chased)]. V(cat) is revised completely as the dot product between the attention weights and the Value Vectors for "mat" and "chased".

V_n(cat) (revised) = W_a_m(cat).V_{mn} (mat, chased)

So, V(cat) (revised) = [0.31, 0.69] . [[0.2, 0.4, 0.5], [0.5, 0.5, 0.8]]
= [0.407, 0.469, 0.707]

This transformation iterates across multiple layers in the Transformer, refining word associations through self-attention, feedforward layers, and backpropagation.

Once all tokens are encoded into vector space, the model generates responses sequentially. Starting with a 'Start' token, it predicts the most probable next token using decoding strategies (e.g., Greedy Search, Beam Search). For example, if the model outputs "The cat was chased by," the probabilities for the next word might be: P(mouse) = 0.004, P(dog) = 0.5, P(man) = 0.05, P(France) = 0.000008. The highest probability word, "dog," is selected, producing the sentence: "The cat was chased by the dog."

ISSN No: 2349-059X
Peer Reviewed & Refereed Journal
National Research Journal

LIMITATIONS OF THE MODEL'S EPISTEMOLOGY:

While such models are extremely efficient and helpful for a variety of tasks, we have to be mindful of their limitations in terms of structural epistemology. Some of them are as follows:

1. **The AI Model as an Agent - The Absence of Qualia and True Comprehension:** To frame the issue epistemologically, we can consider an LLM as an agent interacting with the world. The degree to which it "understands" is thus a function of this interaction. However, an LLM is not conscious—it lacks qualia, the subjective experience of perception, and therefore does not possess true comprehension. LLMs process words through vectorization, mapping statistical associations between tokens, but this is not equivalent to grasping meaning. From an epistemological standpoint, this distinction aligns with Immanuel Kant's differentiation between the noumenon (the thing-in-itself) and the phenomenon (our perception of it). While Kant argued that humans could never experience the noumenon directly, AI does not even engage in phenomenality—it lacks the a priori cognitive structures, or "Categories of Understanding," necessary for human experience. These include Quantity (Unity, Plurality, Totality), Quality (Reality, Negation, Limitation), Relation (Inherence & Subsistence, Causality & Dependence, Community), and Modality (Possibility–Impossibility, Existence–Nonexistence, Necessity–Contingency). An AI model, in principle, cannot possess these categories. For instance, it does not inherently recognize existence or nonexistence: it cannot comprehend that "a dog that once lived and is now dead" has transitioned from existence to nonexistence. Rather than engaging with concepts, it operates within a self-referential loop, interpreting words only through their relationships with other words. This circularity precludes true understanding, distinguishing AI processing from human cognition, which involves embodied experience and non-representational qualia.
2. **The Epistemic Inadequacy of Semantic Clustering and Vectorization:** Even at a technical level, LLMs do not 'understand' words in the way humans do. Rather than grasping concepts, they rely on semantic clustering—identifying statistical proximities among words. For instance, an LLM may associate "cat" and "dog" because they frequently co-occur in similar contexts, but it does not infer that both are four-legged, domesticated animals. Their apparent similarity is based purely on co-occurrence patterns, not an ontological recognition of shared attributes. Furthermore, while LLMs transform words into high-dimensional vector representations, these representations remain abstractions of linguistic data, not reality itself. The notion that meaning can be fully represented through numerical embeddings is flawed; no number of dimensions in a vector space can capture the depth of lived experience. Human cognition, by contrast, relies on qualia, embodied understanding, and contextual depth — elements that go beyond symbolic representation. Michael Polanyi's theory of tacit knowledge further underscores this limitation. He famously asserted, "We can know more than we can tell." That is, much of human knowledge is implicit, inexpressible in words, and fundamentally non-articulable. If we represent total knowledge as A, then the subset that can be articulated as words (B) is only a fraction of it, with C representing the inarticulable, tacit knowledge, i.e. A = B + C. AI does not merely fail to capture C — it does not even understand the information it processes in B. Even in LLMs with thousands of dimensions in their vector embeddings, this limitation persists.

3. **The Mechanism of Token Prediction and the Lack of Original Thought:** AI - generated output is produced through a sequence of probabilistic token predictions. Each word is selected based on statistical likelihood rather than conceptual coherence. This means that, unlike human thought, AI-generated text is not formed holistically as a complete idea before expression. Instead, sentences emerge word-by-word through stochastic modeling. Consequently, LLMs do not generate original thought; they recombine existing human expressions in statistically probable ways. At best, they produce novel permutations of pre-existing arguments, but they do not engage in genuine reasoning or conceptual synthesis. This process, often described as "stochastic parroting," creates the illusion of intelligence while lacking actual cognition.
4. **The Limits of AI in Counterfactual and Causal Reasoning:** AI's reliance on statistical correlations rather than causal relationships further constrains its epistemic capabilities. Counterfactual reasoning — engaging with "what-if" scenarios — requires an understanding of cause and effect (as also one of the Kantian 'categories of understanding' referred to, above), something current deep learning models lack. While efforts have been made to integrate causal models into AI, these approaches contradict the fundamental paradigm of deep learning, which is to allow models to develop their own representations autonomously rather than being explicitly programmed with causal logic. Even state-of-the-art approaches, such as Counterfactual Generative Adversarial Networks (GANITE), struggle with this issue. As Yoon, Jordon, and van der Schaar (2018) state: "only the factual outcome is actually observed (revealed), counterfactual outcomes are not observed and so the entire vector of potential outcomes can never be obtained" *(Yoon et al., 2018, GANITE: Estimation of Individualized Treatment Effects Using Generative Adversarial Nets)*. In other words, while counterfactual GANs can generate plausible alternative outcomes, they lack the structural framework to ensure these outcomes reflect true causal relationships. They rely on statistical proxies rather than an actual causal model of the world. A similar inadequacy of deep learning models in understanding causation has been noted by leading AI researchers, including Yoshua Bengio. "He (Bengio) believes it (Deep Learning) won't realize its full potential, and won't deliver a true AI revolution, until it can go beyond pattern recognition and learn more about cause and effect. In other words, he says, deep learning needs to start asking why things happen." *(Yoshua Bengio interview remarks on causality and AI - 'An AI Pioneer Wants His Algorithms to Understand the 'Why'' (Wired, 2019))*. The introduction of explicit causal models into AI challenges the core statistical nature of neural networks, raising the question of whether deep learning can ever develop true causal understanding without human intervention. Moreover, even if such models are implemented, their scalability remains uncertain—how feasible is it to encode logic for every possible interaction an AI system may encounter?

CASE-BASED EVALUATION IN VUCA CONTEXTS

1. **The Cyc Project - AI's Failure in Common Sense Reasoning:** The Cyc Project (1984) aimed to encode human common sense into a structured AI system but failed after decades of development. Its limitations illustrate AI's struggles in VUCA environments:
   - Volatility: Cyc required explicit rules for every scenario and couldn't adapt dynamically (e.g., rerouting due to a protest blocking a road).

ISSN No: 2349-059X
Peer Reviewed & Refereed Journal
National Research Journal

- Complexity: Unlike humans, who generalize intuitively, Cyc needed every rule articulated (e.g., "Don't put a laptop in a microwave"), making it impractical for real-world decision-making.

2. **Air Canada Chatbot (2023):** Misled a passenger about fare policies, leading to financial loss. The court ruled AI-generated misinformation as the airline's responsibility, highlighting AI's failure in policy-sensitive decisions.
3. **COVID-19 Diagnosis AI (2021):** Machine learning models misidentified patterns due to biases, proving unreliable in clinical use. Without causal reasoning, AI mistakes correlations for insights, failing in high-stakes contexts.

Implications for Sustainable Businesses & Governments in VUCA Contexts:

Despite their epistemic limitations, AI models—particularly LLMs and NLPs—have become indispensable in business automation. Their ability to generate coherent outputs, despite lacking true comprehension, allows them to efficiently handle structured and semi-structured tasks. However, their application in high-stakes decision-making, especially in VUCA (Volatile, Uncertain, Complex, and Ambiguous) environments, demands a more nuanced approach. AI should be leveraged where appropriate, but businesses and governments must address its inherent deficiencies to ensure long-term sustainability.

To navigate this AI-human hybrid landscape, the following strategic imperatives should be considered:

- **Redefine AI's role in decision-making**: AI should augment, not replace, human judgment, particularly in contexts requiring qualitative reasoning, counterfactual analysis, and ethical considerations.
- **Cultivate human expertise in critical, non-automatable domains**: Businesses must actively develop skills in adaptability, contextual awareness, strategic foresight, and creativity—areas where AI falls short.
- **Reorient workforce structures to avoid over-reliance on AI**: Employees should transition away from routine, automatable tasks and develop competencies that complement AI rather than compete with it.
- **Leverage AI for efficiency, but acknowledge its epistemic limitations**: AI is highly effective in structured environments but cannot autonomously navigate ambiguity, causality, or emergent complexity.
- **Adopt policy measures to mitigate AI-driven economic shifts**: Governments should anticipate the societal impact of automation, addressing labor displacement, skill mismatches, and the economic implications of declining demand for routine intellectual labor.
- **Balance AI's benefits with pragmatic expectations**: While fostering creativity and analytical thinking is essential, it is unrealistic to expect universal proficiency in these areas. AI should be a tool that enhances human capability rather than a substitute for human intuition and expertise.

REFERENCES:

1. Kant, I. (1781/1787). *Critique of Pure Reason* (P. Guyer & A. W. Wood, Trans.). Cambridge University Press. (Original work published 1781/1787)

2. Polanyi, M. (1958). *Personal Knowledge: Towards a Post-Critical Philosophy*. University of Chicago Press.
3. Polanyi, M. (1966). *The Tacit Dimension*. University of Chicago Press.
4. Yoon, J., Jordon, J., & van der Schaar, M. (2018). GANITE: Estimation of Individualized Treatment Effects Using Generative Adversarial Nets. In *International Conference on Learning Representations (ICLR)*.
5. Metz, C. (2019, November 21). An AI Pioneer Wants His Algorithms to Understand the 'Why'. *Wired*.
6. Quach, K. (2024, February 15). Air Canada must pay damages after chatbot lies to grieving passenger about discount. *The Register*.
7. Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A., Etmann, C., McCague, C., Weir-McCall, J., Teng, Z, Gkrania-Klotsas, E., Rudd, J., Sala, E., & Schönlieb, C. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence*, *3*(3), 199-217.