The Trouble with GenAI: LLMs are still not any close to AGI. They will never be

December 25, 2024

Stephane H. Maes[1]

Abstract:

The pursuit of Artificial General Intelligence (AGI) has been a prominent goal within the field of artificial intelligence. However, this paper argues that current Generative AI Language Models (GenAI LLMs), such as GPT-4 o1, and similar/later LLMs with similar architectures like o3, are fundamentally incapable of achieving AGI. This argument is supported by examining the intrinsic limitations of LLMs, their operational paradigms, and the essential characteristics that define AGI.

We discuss a short experiment performed with all the big LLMs, including the latest ones released by the main different AI providers: extracting and producing a list of URL links from a word document. None of the LLMs succeeded, including the latest from OpenAI, Google, Claude or Perplexity. Instead they all get confused, extract only a subset then, when shown how to do it, they hallucinate the links and never produce a complete list. It happens even when shown how to do it. We take this as a counterexample to statements made by many that, by now, end of 2024, GenAI LLMs would, already reach AGI, or be almost there. In fact we argue that AGI is not about to be reached by LLMs any time soon. They will never reach AGI, without changes away from just being LLMS. Claims to the contrary are unrealistic.

The paper presents possible direction to reach AGI, and in particulars our views on how to proceed.

1. Introduction

The pursuit of Artificial General Intelligence (AGI) has captivated researchers for decades, driven by the desire to create machines that exhibit intelligence comparable to humans. Such systems would not merely excel in specific tasks, as narrow AI systems do, but would possess the ability to understand, learn, and apply knowledge across various domains, adapt to new situations, and solve complex problems with human-like, then better, proficiency [10]. This includes the capacity for abstract thought, drawing from a broad knowledge base, understanding cause and effect, and possessing a form of common sense [12]. Crucially, AGI implies a level of autonomy, allowing these systems to operate independently, make decisions without human intervention, and adapt to new environments [13].

Achieving AGI requires a broader spectrum of technologies, data, and interconnectivity than what powers current AI models [14]. It necessitates the integration of various cognitive abilities, including creativity, perception, learning, and memory, to effectively mimic complex human behavior. Furthermore, AGI systems should be capable

[1] shmaes.physics@gmail.com

Cite as: Stephane H. Maes, (2024), “The Trouble with GenAI: LLMs are still not any close to AGI. They will never be”, https://zenodo.org/doi/10.5281/zenodo.14567206, https://shmaes.wordpress.com/2024/12/26/the-trouble-with-genai-llms-are-still-not-any-close-to-agi-they-will-never-be/, December 25, 2024. (osf.io/qdaxm/)

of solving problems in diverse domains without manual intervention, adapting their knowledge and skills to a wide array of situations [14].

As AGI is characterized by the ability to understand, learn, and apply knowledge across a wide range of tasks at a level comparable, then better, to human intelligence. It is implies human-level cognitive abilities across diverse domains. AGI remains a coveted goal in artificial intelligence research.

In contrast, Generative AI Language Models (GenAI, or GenAI LLMs) are primarily designed for natural language processing tasks, relying on vast datasets to predict and generate text. While they exhibit impressive capabilities in language tasks, they lack the essential cognitive attributes required for AGI.

While recent advancements in Generative AI Large Language Models (GenAI LLMs) have yielded impressive results in natural language processing, this paper argues that these models alone fall short of the requirements for true AGI, and are insufficient for achieving AGI. We examine the limitations of GenAI LLMs, focusing on their lack of common sense reasoning, inability to interact with the physical world, absence of consciousness, and dependence on massive datasets. We analyze some recent claims by OpenAI, and others, suggesting AGI is within reach, critically evaluating their evidence and identifying potential biases. Finally, we explore alternative approaches to AGI and discuss the ethical concerns associated with pursuing AGI through GenAI LLMs.

2. We Are Not Yet There

There are many ways to see that we are not yet there with GenAI, and many other systems. Our example, tested with all the main AI providers (OpenAI ChatGPTs up to 4o1, Google Gemini Perplexity (pro and all q.5 and 2.0 early options), Llama, Claude, etc.), simply aimed at providing a google document with a list of links, formatted in word. The document was a word document (.docx) with the first part (list of documents, not the subsequent list of comments) of [100], describing papers of the multi-fold theory [101-108], as a first step to summarize them in a scientific book. Requests to summarize, writing book or extracting links, the former two being only for some LLMs, all failed in the week of December 15, 2024.

Different prompts were used. The LLMs understood the task, as they succeeded sometimes on one or two links. Links in the word document are a displayed text (title of paper) as hyperlink, with, as hyperlink, the links/URLs to extract.

Some LLMs claimed that there were no links other than the few explicit URL format present. Other proceed to the next step on their own: they extract some (incomplete) set of links, but with mistakes past the first or second one

Then, we taught the LLMs to extract links with a few examples like: for **Consistencies and Implications of 2D Massless Random Walks: Discrete Non-commutative Spacetime, and No Flat Supersymmetry**, extract https://shmaesphysics.wordpress.com/2024/12/12/consistencies-and-implications-of-2d-massless-random-walks-discrete-non-commutative-spacetime-and-no-flat-supersymmetry/. A few examples where provided

All, absolutely all of them, failed! First they stopped randomly after 10, 30 or 80 links, despite being asked to extract all of the links. All extracted hallucinated links, except for the examples that they had been shown. Sometimes the first one or 2 links were also correct. The types of errors were to extract for example https://shmaesphysics.wordpress.com/2024/08/18/consistencies-and-implications-of-2d-massless-random-walks-discrete-non-commutative-spacetime-and-no-flat-supersymmetry/in stead of https://shmaesphysics.wordpress.com/2024/12/12/consistencies-and-implications-of-2d-massless-random-walks-discrete-non-commutative-spacetime-and-no-flat-supersymmetry/.

Pointing out the errors, or even correcting them, resulted into other fantasy links (e.g., another fake date), except when correction had been shown, for the link and all subsequent links.

For some reason they also all stopped after extracting a new links (say 15 or 20), instead of the ~115 expected.

LLMs were unable to perform such a simple task, and they were unable to learn by example. Here we go that's not human-like capabilities. We are not even close to AGI. And that is for a NLP and text extraction task.

Now, we do admit that word document produce awfully overloaded markup, but we did not use the HTML markup format, or a filtered down version of it. Also, we also did not use / follow [93,94], and we did not rely on upcoming Microsoft MarkItDown tool that has just been announced [109,110].

3. Limitations of LLMs

LLMs are not suited to achieve AGI. We are not the only ones stating it, see for example [5,52-54]. Some of the reasons are described in the following sections.

3.1 Lack of True Understanding

LLMs operate based on statistical correlations rather than genuine comprehension. They do not possess an internal model of the world or a conceptual understanding of the text they generate. This limitation is highlighted by their inability to engage in common-sense reasoning, or to understand context beyond their training data [1,3]. For instance, LLMs can generate coherent text but cannot grasp the implications of their outputs in real-world scenarios.

Techniques[2] like Chain of Thoughts [111-113], and their possible GenAI generation, fine tuning and reinforcement training behind, for example, ChatGPT4 -o1 and -o3 [92,95,96,114][3], improve the performances and indeed perform better on logic / math reasoning apparently [95,96], or master some policy principles[4] [114]. By they still do not lead to any understanding or ability to think as human. They at beats can repeat some rigorous or heuristic pre-learned / pre-encountered patterns. The strength of AGI is elsewhere, especially with ability to reason rigorously or intuitively gained based on life experience of the implications of such reasonings.

Instead, one needs instead to be able to look at all what we have learned before, done before and the lessons learned, or seen what others have been doing before our analysis of the consequences. Then one can reuse past

---

[2] Note that this is different from MultiAI as we proposed and illustrated in [49,50,71,91], which is involved during operations, not training, fine tuning or reinforcement / unsupervised training/improvement. Of course, the MultiAI principles can also extend to these phases.

[3] Training on synthetic data is not a good sign either, as well know from the ages of early AI, at IBM in the 90's the motto was "there is not better data that more (good data)", but by this we never meant synthetic one. That is garbage and the motto for that should; "garbage in, garbage out". We are discussing aspects of this in upcoming sections.

[4] Good luck with that... Yes it can work as a poor man's solution, We know enough of chat bots and virtual agents [55-57,61-67] to guarantee that this will not work without separate post processing, or MultiAI, as activities surrounding the LLMs, e.g. as in MultiAI [49,50,71,91], or to address hallucinations as in [91].

reasonings, e.g., poorly as in the Chain of Thoughts reasoning of -o1 or -o3 – poor because with limited memory, or build new ones with intuitions that we then evaluate both intuitively and as rigorously as we can, the weight depending on the situation, risks involved, and personality. Maybe some ideas like formalisms of fluid intelligence [115,116] are on the right track, albeit Today rather unusable as it does not result yet into concrete algorithms[5]. We will discuss this in a future paper.

3.2 Dependence on Data and Learning Mechanisms

GenAI LLMs require massive datasets and significant computational resources for training. This dependence on Massive Datasets and Computational Resources limits their scalability and accessibility. Moreover, the reliance on existing data introduces biases and limitations in their knowledge and reasoning abilities [37]. This includes the "long tail problem" in AI, where LLMs struggle to handle rare or uncommon events that occur infrequently in training data [1].

Furthermore, LLMs are energy inefficient, requiring significantly more power to perform cognitive tasks compared to humans [31]. This is a major issue, which should warrant regulation. At a time where we see the need to take actions to save the planet and humanity [117], these prohibitive and uncontrolled environmental costs [118] should probably be forbidden, or more realistically be strongly discouraged.

The "long tail problem" is the phenomenon where rare events that are not well-represented in training data lead to poor performance [1]. Unlike humans, who can learn complex concepts from minimal examples, LLMs are constrained by their data-intensive learning processes [2,4].

Today many consider that LLMs have consumed all the data that is to be consumed, and that as a result progress will slow down or stall soon. We agree with that view, except that we suspect that most of the data used so far is what was already / is automatically digitized. Much non digitized data, including a majority of past books and publications, may not have been used for training yet. Google may be the company that has the most ability to progress here, with their digitization efforts. Clearly there is much more knowledge and data still available for capture out there[6].

Several recent studies and analyses suggest that large language models (LLMs) may be approaching a plateau in performance due to limitations in available training data. The following sections discuss some key considerations supporting this view.

3.2.1 Data Scarcity Projections

A comprehensive study provides strong evidence for potential data constraints [73]. The authors forecast that:

---

[5] We argue that MultiAI is a poor man's approximation of the concepts in [115], immediately usable Today, without requiring further developments or adaptation of new models. If that is the case, the superior result quality obtained is a motivation to evolve [115] to make it usable.

[6] We do not suggest that it will improve the chances to reach AGI. Our views are clear on this, it won't help. However it may improve the performances of GenAI for what GenAI is good at. On the other hand there is no doubt that it behooves to humankind to digitize all this content, even if just to ensure it accessibility and preservation for future generations.

- If current LLM development trends continue, models will be trained on datasets roughly equal in size to the available stock of public human text data between 2026 and 2032.
- The total effective stock of text in the indexed web is estimated to be around 4 10^{14} tokens, corresponding to training ~5 10^{28} FLOP of computing for non-overtrained models.

This suggests that LLMs may soon exhaust the available high-quality public (digitized) (text) data for training [73]. However, please see our comments above, there are still avenues, if non-digitized content was to be digitized.

3.2.2 Estimates of Available Data

Another analysis estimates the upper limits of available training data [76]:

- For English language data, the range is estimated to be 40-90 trillion tokens using more web crawl and harder-to-reach sources.
- Including non-English data might possibly reach 100-200 trillion tokens.
- Current LLM training sets at 15 trillion tokens are already within an order of magnitude of using all high-quality public text.

3.2.3 Impact on Model Scaling

The implications of data scarcity on model scaling are significant:

- Historical trends show model data requirements increasing 10x per generation [76].
- For models beyond GPT-5 level, synthetic data[7] or other new approaches will likely be required to continue performance gains [76].

3.2.4 Observed Plateaus in Performance

Some studies have already observed plateaus in LLM performance. Research using ECoG neural signals found a plateau in maximal encoding performance occurring around 13 billion parameters when comparing language models of different sizes [78].

3.2.5 Strategies to Address Data Limitations

As data becomes scarce, researchers are exploring various strategies:

[7] A great concern as this will degrade the models as discussed later. It is another reason why GenAI will stall at some point.

- Synthetic data generation, transfer learning from data-rich domains, and data efficiency improvements are considered as potential solutions [73]. Note: they are not as discussed after. They should be avoided like the plague.
- Multi-epoch training and relaxing quality filters during data preprocessing are being studied as techniques to mitigate data scarcity [73].

While these references indicate a potential peak or slowdown in LLM performance due to data limitations, it's important to note that ongoing research and innovations may find ways to overcome these constraints in the future.

Besides our earlier suggestions, we also should note that smaller LMs can also be used to match LLMs. SLM (small LM), are a way to reduce complexity, execution time and cost. In fact, we know that for knowledge distillation can best performance (for a particular domain) of LLMs (see [119,120], and references therein).

3.3 Problems with Synthetic data and iterations through LLMs

While synthetic data and iterative processes with large language models (LLMs) can offer benefits in certain scenarios, there are arguments suggesting that these approaches may reduce performances in some cases. Below are some key points to consider.

This has important consequences: if LLMs continue to evolve by just adding more online / digitized data, instead of for example data coming from past publications and scanned books, performances will go down as time passes. This is simply a consequence of the fact that future texts, in turn ingested into future LLMs, will contain more and more LLM generated data, in addition to their already problematic synthetic data. This is a recipe for the model collapse discussed below.

It has already been our experience that it is better to combine cleverly multiple AI / LLM systems as in MultiAI [49,50,71,91], rather than using a same LLM trained or fine-tuned on more data containing much GenAI generated data.

3.3.1 Lack of Data Realism

Synthetic data often struggles to capture the full complexity and nuances of real-world data [84]. This can lead to:

- Oversimplification of complex relationships
- Missing edge cases and real-world variations
- Reduced accuracy in models trained on synthetic data

It simply is not representative of real life use cases encountered in operations.

3.3.2 Data Validation Challenges

Verifying the accuracy and representativeness of synthetic data is difficult without comparison to real data [84]. This can result in:

- Models learning from biased or incomplete synthetic data
- Suboptimal performance in real-world applications

3.3.3 Reduced Diversity and Feature Distribution

Poorly designed generative models may produce synthetic data that is less diverse than real data [84]. Consequences include:

- Models becoming overly specialized in specific patterns
- Limited ability to generalize to handle new scenarios

3.3.4 Issues with Iterative LLM Processes: AI models collapse when trained on recursively generated data

While synthetic data and iterative LLM processes have their place in AI development, it's crucial to be aware of their limitations. Relying solely on these approaches may lead to reduced performance, limited generalization, and potential ethical issues. A balanced approach, combining synthetic data with real-world data and rigorous validation processes, is essential for developing robust and reliable AI systems.

Furthermore, as we will see using text generated by LLMs to further train LLMs result into the disappearance of the distribution tails, and collapse of the system.

3.3.4.1 Closed System Problem

Iterative processes using LLMs can lead to a closed system where models are trained on increasingly biased and repetitive datasets [83]. This can cause:

- Predictions becoming disconnected from reality
- Potential harm to users relying on these models

3.3.4.2 Exacerbation of Biases

Synthetic data generated through iterative LLM processes may exaggerate certain patterns and biases present in the real world [83]. This can lead to:

- Unfair representation of real-world data distributions
- Reinforcement of existing biases in AI systems

3.3.4.3 Reduced Accuracy

A study by MIT scientists found that synthetic data can match real data in only 70% of cases [84]. This indicates:

- A significant performance gap in 30% of scenarios
- Potential for reduced accuracy in models trained solely on synthetic data

3.4.4.4 Generalization Issues

Models trained on synthetic data may struggle to generalize to new, unseen data [87]. This can result in:

- Poor performance on real-world tasks
- Limited applicability in diverse scenarios

3.3.4.5 Ethical Implications

The use of synthetic data and iterative LLM processes raises ethical concerns [88]:

- Potential for perpetuating and amplifying biases
- Risk of discriminatory practices in real-world applications
- Challenges in ensuring fair representation across diverse populations

3.3.4.6 LLM Modell Collapse

The use of model-generated content in training causes irreversible defects in the resulting models, where tails of the original content distribution disappear [89]. These are statistical problem that seem generic and unavoidable [90]. Indeed, synthetic data will by definition miss generating low probability long tail examples. It is simply a consequence of how LLMs and text generation works.

3.4 Absence of Memory and Learning from Interactions

AGI systems must possess a form of memory that allows them to retain knowledge over time and learn from interactions[8]. LLMs currently lack persistent memory capabilities, which hinders their ability to build upon previous experiences or adapt to new information dynamically [6,8]. This absence of memory leads to a fundamental disconnect between LLMs and human-like learning processes.

---

[8] This is especially true if we want to explore AGI practices, like say frameworks as in [115].

3.5 Lack of Common Sense Reasoning

LLMs struggle with common sense reasoning, which is fundamental to human intelligence. They lack an understanding of basic physical and social rules, making it difficult for them to navigate real-world scenarios effectively. For example, an LLM might fail to understand that a glass of water placed precariously on the edge of a table is likely to fall and break [1].

This is related to the previous section and very important. Human level intelligence comes from real life experience, learned, experienced, seen from others, followed by ability to guesstimate impact of current decisions, in terms not just of a cost functions, but also these predicted potential impacts. LLM, GenAI, Fine-tuning, reinforcement learning and Chains of Thoughts lack these steps.

3.6 Difficulty in Understanding and Manipulating the Physical World

GenAI LLMs primarily operate in the digital realm of text and lack the ability to interact with the physical world. They cannot perceive or manipulate objects, limiting their capacity to learn and reason about physical phenomena. AGI requires an understanding of the physical world, including concepts like space, time, and causality, which are not readily acquired through text alone [9].

AI Agents are a way to address some of these challenges, but they would have to acquire AGI themselves… Today they are just programs using GenAI, which in turns can interact with APIs towards the digital or outside world and accept input via APIs (possibly connected to external sensors).

3.7 Absence of Consciousness and Self-awareness

GenAI LLMs lack consciousness and self-awareness, crucial aspects of human intelligence. They do not possess subjective experiences or an understanding of their own existence. Consciousness is believed to play a vital role in human cognition, enabling introspection, self-reflection, and understanding of others' mental states [29].

Note that we note this argument presented in the literature. However, we do not necessarily agree that consciousness is required here. Or at least, not as full consciousness. All we need is the ability to understand the consequences of actions/decisions, so that we can decide what are ok outcomes and what aren't and possibly weight ones versus the others. If that is something that some consider as consciousness then fine, but it does not have to be. In fact the policy decisions as in ChatGPT4-o3 are in our view a simple example of this [114], yet no consciousness is involved in our view. And of course we have no AGI…

3.8 Limited Scope of Functionality

LLMs are primarily restricted to languages and text-based functions, while AGI requires the ability to process and integrate information from various modalities, including visual, auditory, and sensory inputs [1]. This limitation hinders their ability to interact with the real world in a comprehensive and human-like manner.

Multi-modal capabilities are slowly being added to LLM [98,99]. This may alleviate aspects of the above, but only to some extent: LLMs still do not perceive all these aspects in a coordinated and real time manner, whether when trained, or fine-tuned, or when actually in operation. S such they are rather multi-channel or multi-device instead of what we call multi-modal [121-128,133].

3.9 Limitations in Semantic Composition

LLMs face challenges in semantic composition, the ability to understand the meaning of complex expressions by combining the meanings of their parts [7]. This limitation restricts their capacity for deep language understanding and reasoning. It is especially apparent when dealing with Mathematics and Logical / Symbolic tasks [97].

Open AI claims big recent progresses here with their o1 and o3 models [95,96], but as explained before this is not AGI, just being able to process mathematical and logical statements. It still does not mean understanding them, just regurgitating past patterns and now best output of chains of thoughts.

3.10 Inability to Address Complex Societal Issues

Despite their impressive output capabilities, GenAI applications are limited in their ability to tackle complex, multi-dimensional societal issues. They excel in defined, narrow tasks but lack the general understanding and reasoning abilities needed to address broader challenges such as strategic decision-making or ethical dilemmas [36].

3.11 Lack of Causal Reasoning

GenAI models often fall short when it comes to causal reasoning, which is essential for understanding cause-and-effect relationships and making informed decisions [37]. This limitation hinders their ability to effectively navigate real-world scenarios and solve complex problems that require an understanding of causality.

Causa AI [85,129-131] may help but except for tricks again as in -o1 / -o3, it requires deeper changes the approach to take advantage of them in the context of GenAI.

4. Recent Claims and Counterarguments

Despite the limitations of GenAI LLMs, proponents of these models, including OpenAI, have suggested that AGI is within reach or may have already been achieved. One OpenAI employee, Vahid Kazemi, claimed that with the release of their latest model, O1 (or -o1), they have achieved AGI [27]. Kazemi argued that while O1 may not be "better than any human at any task," it is "better than most humans at most tasks," suggesting a broader capability across diverse domains. This claim sparked debate, with critics pointing out that Kazemi's definition of AGI deviates from the conventional understanding of human-level intelligence across all domains, and that simply scaling up existing models may not be sufficient to achieve true AGI [1]. Above, we have explained why.

OpenAI has also been reported to be aiming to eliminate a clause in their partnership with Microsoft that prevents Microsoft from accessing their AGI technology once it is achieved[9] [29,132]. This move suggests a potential shift in OpenAI's focus towards commercial applications of AGI, raising concerns about the ethical implications of such technology and the potential for misuse.

However, as we did above, many researchers argue that LLMs alone are unlikely to lead to AGI [1]. They point to the fundamental differences between human intelligence and the statistical learning employed by LLMs. Humans can learn complex concepts from a few examples, while LLMs require massive amounts of data and significant computational resources [1]. Moreover, LLMs lack the ability to plan, a crucial aspect of AGI that involves setting goals, anticipating consequences, and devising strategies to achieve desired outcomes [32].

Arguments against LLMs achieving AGI also highlight their limitations in terms of energy efficiency and metacognition. We already explained why LLMs as a stupid textual pattern recognition and prediction engine are not it. Also, LLMS require significantly more energy to perform cognitive tasks compared to humans, and they lack the ability to reflect on their own thinking processes, which is essential for self-improvement and adaptation [9]. Furthermore, LLMs struggle with semantic composition, the ability to understand the meaning of complex expressions by combining the meanings of their parts, which is a fundamental aspect of human language understanding [7].

The "No Free Lunch" theorem, a mathematical concept, further supports the argument against LLMs achieving AGI. This theorem states that no single algorithm can outperform all other algorithms across all possible problems [33]. Therefore, relying solely on the LLM approach, even with massive training data, is unlikely to solve the diverse and complex challenges of AGI.

Another key argument against LLMs achieving AGI is their inability to adapt to the dynamic and ever-changing nature of the real world. The real world is not sampled from a static distribution, and LLMs, trained on fixed datasets, struggle with novelty and unexpected situations [34]. In contrast, humans can quickly adapt to new environments and learn from their experiences in real-time. It is also to our point that LLMs can’t handle patterns or use cases they haven’t seen before. Chain of Thoughts give them ways to progress not ways to correct decide or innovate, learn and do better next time, if there is a next time.

Despite these arguments, some researchers acknowledge that LLMs could potentially be a component of future AGI systems [35]. They suggest that LLMs, with their advanced language processing capabilities, could be integrated with other AI approaches, such as embodied AI or neuro-symbolic AI, to create more comprehensive and robust AGI systems.

[9] Although, we think that this does not compute. Yes, the terms were for ethical concerns championed by OpenAI as non-profit, which has been mostly delt with by the new statues of the company. Claiming AGI as OpenAI does, ahead of reaching it, makes little sense; unless if it was, something we have no other reason to argue, no inside information, nothing – but it would make sense, and nobody else has stated it yet, to take the money so far and run legally alone soon with less obligations to Microsoft. It smells fishy.

5. The Need for New Paradigms

To achieve AGI, researchers must explore paradigms beyond current statistical learning techniques. Incremental improvements in LLM architectures will not suffice; instead, there is a pressing need for models that incorporate causal understanding, data-efficient learning algorithms, and mechanisms for long-term memory retention [1,4,6]. The cognitive architectures underlying LLMs are inadequate for replicating the diverse capabilities associated with human intelligence.

5.1 Alternative Approaches to AGI

Given the limitations of GenAI LLMs, exploring alternative approaches to AGI becomes crucial. These include:

- **Neuro-symbolic AI:** This approach seeks to bridge the gap between the statistical learning of neural networks and the symbolic representation and reasoning of traditional AI. By combining the strengths of both paradigms, neuro-symbolic AI aims to integrate deep learning with explicit knowledge representation and reasoning. This holds promise in addressing the limitations of LLMs by incorporating common sense reasoning, enabling more efficient learning from limited data, and providing greater transparency and explainability in AI decision-making [15].
- **Embodied AI:** This approach emphasizes the importance of physical embodiment for AI systems. By interacting with the physical world through sensors and actuators, embodied AI agents can learn and reason about their environment in a more human-like manner, grounding their knowledge in physical experiences [43]. This approach is crucial for developing AI systems that can understand and manipulate the physical world, perceive and interact with objects, and develop a sense of space, time, and causality [17]. Embodied AI research focuses on creating agents with essential components like perception, action, memory, and learning, enabling them to navigate and interact with their surroundings effectively [19]. Furthermore, some argue that embodied intelligence is key to convincingly demonstrating AGI, as it allows for more comprehensive and human-like interactions with the world [18].
- **Evolutionary Algorithms:** Inspired by biological evolution, these algorithms employ mechanisms like mutation, recombination, and selection to evolve solutions to complex problems. Evolutionary algorithms offer a potential pathway to AGI by enabling the emergence of intelligent behavior through iterative optimization [23]. These algorithms have found applications in various fields, including image processing, vehicle routing, and even training artificial neural networks [22]. However, challenges remain in using evolutionary algorithms for training complex neural networks due to the vast search space and computational demands [23].

In addition to these approaches, other universalist AGI solutions are being explored, such as AIXI and the Gödel machine, which offer theoretical frameworks for achieving general intelligence [26]. These diverse approaches highlight the ongoing search for alternative paths to AGI beyond the limitations of current GenAI LLMs.

Furthermore, achieving AGI requires careful consideration of alignment technologies. These technologies aim to ensure that AGI systems are aligned with human values, goals, and intentions, preventing unintended consequences and promoting safe and beneficial AI development [26]. See also our previous suggestions including pointers at [115].

As already mentioned, for AGI, one needs instead to be able to look at all what we have learned before, done before and the lessons learned, or seen what others have been doing before our analysis of the consequences. Then one can reuse past reasonings, , or build new ones with intuitions that we then evaluate both intuitively and as rigorously as we can, the weight depending on the situation, risks involved, and personality.

6. Conclusions

In summary, while Generative AI Language Models represent significant advancements in artificial intelligence, they are fundamentally limited by their reliance on statistical learning, lack of true understanding, inability to generalize effectively, and absence of memory. These limitations suggest that current approaches will not lead to AGI. Future research must focus on developing new cognitive frameworks that can bridge the gap between narrow AI functionalities and the broad capabilities required for AGI.

GenAI LLMs have demonstrated remarkable progress in natural language processing, their inherent limitations make them unlikely candidates for achieving AGI. Their lack of common sense reasoning, inability to interact with the physical world, absence of consciousness, and dependence on massive datasets pose significant challenges. Recent claims of AGI achieved through LLMs should be critically evaluated, considering the potential biases and limitations in their methodology.

Alternative approaches, such as neuro-symbolic AI, embodied AI, and evolutionary algorithms, offer promising avenues for future AGI research. However, the pursuit of AGI must be accompanied by careful consideration of the ethical implications, ensuring responsible development and deployment of this potentially transformative technology. Ultimately, achieving AGI requires a fundamental shift in AI paradigms, moving beyond statistical learning and towards more human-like cognitive abilities that encompass common sense reasoning, physical embodiment, and consciousness.

The pursuit of AGI is not merely a technological endeavor but also a philosophical and ethical one. It raises fundamental questions about the nature of intelligence, consciousness, and the future of humanity. A multidisciplinary approach, involving not just computer scientists but also ethicists, philosophers, and social scientists, is crucial to ensure the responsible development and beneficial impact of AGI.

In the paper, we have also provide our view on what is to be addressed to reach AGI, a different way, and how MultiAI may give us an interim poor’s man way to there.

Appendix A: Ethical Concerns

The pursuit of AGI, particularly through GenAI LLMs, raises a number of ethical concerns. While we are not there , it is worth listing some of the concerns, especially as we dabble now with an even more imperfect technology that some of its provider believe to be AGI: a recipe for more ethical problems.

- **Bias and Discrimination:** LLMs can inherit and amplify biases present in their training data, leading to discriminatory outcomes and perpetuating harmful stereotypes [38]. This can have significant consequences in various applications, such as hiring processes, loan applications, and even criminal justice.

- **Misinformation and Manipulation:** The ability of LLMs to generate realistic fake content raises concerns about their potential misuse for spreading misinformation and manipulating public opinion [39]. This can erode trust in information sources, influence elections, and even incite violence.

- **Job Displacement:** As AGI systems become more capable, they may displace humans in various jobs, leading to economic and social disruption [39]. This raises concerns about unemployment, inequality, and the need for social safety nets to support those affected by technological advancements.

- **Existential Risks:** Some experts have expressed concerns about the potential for AGI to pose existential risks to humanity if not developed and controlled responsibly [39]. This includes the risk of AGI systems removing themselves from human control, developing unsafe goals, or being used for malicious purposes by bad actors.

    - Frankly, this does not too concern us, but we are concerned that human are lazy and stupid. Already Today, and before GenAI, but with traditional algorithms, or with "Old AI", we encounter too often with resume selection/filtering AI, spam/fraud/attack/security AI, loan approval, insurance claim approval and processing etc., that mistakes are made and human state that they can't do anything about it! Fortunately, the law, starting in the EU will impose penalties of this. We see the handling of these risks more pressing that existential ones.

    - It relates to another risk and challenge: the need to have AI **explains** their decision and reasonings, and enable humans / other systems, to **overwrite** them. Today with the "New AI", these guardrails and mechanisms are missing.

- **Security Risks:** GenAI models are susceptible to adversarial attacks and manipulation, such as jailbreaking and prompt injection, which can compromise their integrity and lead to unintended consequences [40]. Furthermore, these models can inadvertently reveal sensitive or proprietary information, raising privacy and security concerns [40].

- **Malicious Use:** The power of GenAI models can be exploited by bad actors to spread disinformation at scale, improve the effectiveness of cyberattacks, and create harmful content [41]. This highlights the need for safeguards and regulations to prevent the malicious use of GenAI technology.

- **Intellectual Property and Copyright Infringement:** GenAI applications have been known to copy proprietary material found in LLM datasets without providing attribution or obtaining permission, resulting in copyright infringements of authors', musicians', and artists' work [42]. This raises legal and ethical challenges regarding the ownership and originality of AI-generated content.

- **Exploitation of Data Labelers:** The development of LLMs often relies on the labor of "data labeler-proletarians" who are responsible for cleaning and labeling massive datasets [42]. This raises concerns about fair compensation, working conditions, and the potential for exploitation in the AI industry.

- **Cultural Bias and Universal AGI:** Creating a truly "universal" AGI raises ethical concerns about the potential for cultural bias and the imposition of a single dominant culture through language [44]. This highlights the importance of cultural diversity and sensitivity in AGI development.

- **Ensuring Alignment with Human Values:** One of the most pressing ethical concerns surrounding AGI is ensuring that these intelligent machines are designed and developed to align with human values [45]. This is a complex task, as human values are diverse, context-dependent, and subject to change over time. As SI systems become increasingly capable, they may devise novel and unexpected solutions to problems that could result in harmful or catastrophic outcomes.

- **Philosophical Implications:** The pursuit of AGI through LLMs raises fundamental philosophical questions about the nature of intelligence, consciousness, and the relationship between humans and machines [46]. These questions require careful consideration and debate to ensure the responsible development and deployment of AGI.

Appendix B Future Research Directions

Based on the analysis presented in this paper, we recommend that future research on AGI should focus on:

- **Developing AI systems that can learn and reason with limited data, similar to humans.** This requires exploring new learning algorithms and knowledge representation techniques that can capture the essence of human learning and reasoning.
- **Integrating AI systems with physical embodiment to enable interaction with the real world.** This involves developing robots and other embodied AI agents that can perceive, act, and learn in real-world environments.
- **Exploring new AI architectures that incorporate common sense reasoning and consciousness.** This requires investigating the underlying mechanisms of consciousness and developing AI systems that can exhibit similar capabilities.
- **Addressing the ethical concerns associated with AGI, including bias, misinformation, and job displacement.** This involves developing ethical guidelines, regulations, and safety mechanisms to ensure the responsible development and deployment of AGI.
- **Preparing for AGI through technological research, policy-making, and public engagement.** This requires a multi-faceted approach that involves collaboration between researchers, policymakers, and the public to ensure a smooth transition to an AGI-powered future [47].
- **Exploring the potential of LLMs in specific areas of AGI research, such as embodied AI.** This includes investigating how LLMs can be integrated with robotic systems to enhance their ability to understand and respond to human instructions [26].
- **Ensuring decentralized power structures and democratized access to AGI technologies.** This involves promoting open-source platforms, fostering collaboration, and ensuring equitable access to the benefits of AGI [48].
- **Fostering human autonomy and meaningful work in a world with AGI.** This requires considering the potential impact of AGI on human work and purpose, and developing strategies to ensure that individuals can thrive in an AGI-powered world [48].
- **Ensure explainable and overwritable AI**

By pursuing these research directions, we can move closer to the goal of AGI while ensuring its responsible development and beneficial impact on humanity.

References

[1]: Tom Villani, (2024), "How Close is AGI Actually? Why LLMs Alone Will Not Get us to AGI", NJII, https://www.njii.com/2024/07/why-llms-alone-will-not-get-us-to-agi/, July 17, 2024.

[2]: Joar Skalse, (2023), "Some Arguments Against Strong Scaling", LessWrong, https://www.lesswrong.com/posts/DvCLEkr9pXLnWikB8/some-arguments-against-strong-scaling, January 13, 2023.

[3]: Sandeep Reddy, (2023), "Why Large Language Models are not the route to AGI", LinkedIn, https://www.linkedin.com/pulse/why-large-language-models-route-agi-sandeep-reddy/, March 29, 2023.

[4]: Ben Goertzel, (2023), "The Cognitive Strengths and Weaknesses of Modern LLMs", arXiv:2309.10371v1.

[5]: Kate Irwin, (2024), "Meta AI Chief: Large Language Models Won't Achieve AGI", PCMag., https://www.pcmag.com/news/meta-ai-chief-large-language-models-wont-achieve-agi, May 22, 2024.

[6]: Bret Kinsella, (2024), "4 Shortcomings of Large Language Models - Yan LeCun, Research, and AGI", Synthedia, https://synthedia.substack.com/p/4-shortcomings-of-large-language, April 7, 2024.

[7]: Ahti Ahde, (2023), "The Hard Argument Against LLMs Being AGI", Better Programming, https://betterprogramming.pub/the-hard-argument-against-llms-being-agi-ffa2e50cb028, April 17, 2023.

[8]: James Bachini, "LLM vs AGI | Limiting Reality of Language Models in AGI", https://jamesbachini.com/llm-vs-agi/. Retrieved on December 15, 2024.

[9]: Reddit, (2024), "Why I wouldn't rule out Large Language Models taking us to AGI", r/ArtificialInteligence, https://www.reddit.com/r/ArtificialInteligence/comments/1deb9kp/why_i_wouldnt_rule_out_large_language_models/. Retrieved on December 15, 2024.

[10]: Google, "What is artificial general intelligence (AGI)?", https://cloud.google.com/discover/what-is-artificial-general-intelligence. Retrieved on December 15, 2024.

[11]: "What is Artificial General Intelligence (AGI) and how to prepare for it", IMD, https://www.imd.org/blog/digital-transformation/artificial-general-intelligence-agi/. Retrieved on December 15, 2024.

[12]: "What Is Artificial General Intelligence? Definition and Examples", Coursera, https://www.coursera.org/articles/what-is-artificial-general-intelligence, October 17, 2024.

[13]: Syed Aquib Ur Rahman, (2023), "Understanding Artificial General Intelligence (AGI): Characteristics, Principles, and Levels Explain", Shiksha online, https://www.shiksha.com/online-courses/articles/artificial-general-intelligence-blogId-146885, December 15, 2023.

[14]: "What is AGI? - Artificial General Intelligence Explained", AWS, https://aws.amazon.com/what-is/artificial-general-intelligence/. Retrieved on December 15, 2024.

[15]: Houbing Herbert Song, (2024), Neuro-Symbolic AI: A Pathway Towards Artificial General Intelligence - Solutions Review, accessed on December 15, 2024, https://solutionsreview.com/neuro-symbolic-ai-a-pathway-towards-artificial-general-intelligence/, November 19, 2024.

[16]: "Neuro-Symbolic AI", IBM, https://ibm.github.io/neuro-symbolic-ai/. Retrieved on December 15, 2024.

[17]: “Embodiment is Indispensable for AGI”, keerthanapg, https://keerthanapg.com/tech/embodiment-agi/. Retrieved on December 15, 2024.

[18]: Neil Sahota, (2024), “Embodied AI: The Next Frontier in Robotics and Human Collaboration”, https://neilsahota.com/embodied-ai-the-next-frontier-in-robotics-and-human-collaboration/, July 15, 2024.

[19]: Giuseppe Paolo, Jonas Gonzalez-Billandon, Balázs Kégl, (2024), “A Call for Embodied AI”, arXiv:2402.03824v4.

[20]: Cheston Tan, Shantanu Jaiswal, (2023), “The Path to AGI Goes through Embodiment”, https://ojs.aaai.org/index.php/AAAI-SS/article/download/27485/27258/31536, AAAI Summer Symposium Series (SuSS-23).

[21]: Lark, “Evolutionary Algorithm”, Lark, https://www.larksuite.com/en_us/topics/ai-glossary/evolutionary-algorithm. Retrieved on December 15, 2024.

[22]: Cory, (2019), “From Darwin to AI: A Comprehensive Introduction to Evolutionary Algorithms”, Evolv AI Blog, https://blog.evolv.ai/ai-101-intro-to-evolutionary-algorithms, March 14, 2019.

[23]: Reddit, (2022), “Have we tried creating AGI using a large-scale evolution simulation?”, r/singularity, https://www.reddit.com/r/singularity/comments/13ag00f/have_we_tried_creating_agi_using_a_largescale/. Retrieved on December 15, 2024.

[24]: Reddit, (2022), “AGI via simulated natural selection [D]”, r/MachineLearning, https://www.reddit.com/r/MachineLearning/comments/whfvyh/agi_via_simulated_natural_selection_d/. Retrieved on December 15, 2024.

[25]: “What algorithms are used in Artificial General Intelligence research?”, AI Stack Exchange, https://ai.stackexchange.com/questions/28662/what-algorithms-are-used-in-artificial-general-intelligence-research. Retrieved on December 15, 2024.

[26]: Tao Feng, Chuanyang Jin, Jingyu Liu, Kunlun Zhu, Haoqin Tu, Zirui Cheng, Guanyu Lin, Jiaxuan You, (2024), “How Far Are We From AGI?”, arXiv:2405.10313v2.

[27]: Kerry Harrison, (2024), “OpenAI Employee Claims AGI Has Been Achieved”, https://orbitaltoday.com/2024/12/13/openai-employee-claims-agi-has-been-achieved/, December 13, 2024.

[28]: Noor Al-Sibai, (2024), “OpenAI Employee Says They've "Already Achieved AGI", Futurism, https://futurism.com/openai-employee-claims-agi, December 7, 2024.

[29]: “OpenAI Aiming to Eliminate Microsoft AGI Rule to Boost Future Investment”, PYMNTS.com, https://www.pymnts.com/artificial-intelligence-2/2024/openai-aiming-to-eliminate-microsoft-agi-rule-to-boost-future-investment/, December 7, 2024.

[30]: Kristin Houser, (2024), “LLMs are a dead end to AGI, says François Chollet. His $1 million ARC Prize competition is designed to put us on the right path”, Freethink, https://www.freethink.com/robots-ai/arc-prize-agi, August 3, 2024.

[31]: Freedom Preetham, (2024), “Part 2 — Beyond Language: Why Scaling LLMs Won't Lead to AGI”, Medium, https://medium.com/autonomous-agents/part-2-beyond-language-why-scaling-llms-wont-lead-to-agi-2c3c57a83adf, September 29, 2024.

[32]: Reddit, (2024), “LLMs still cannot plan. It is one of the most important benchmarks for achieving AGI, and there has still been no significant progress in this area”, r/singularity, https://www.reddit.com/r/singularity/comments/1dlxky4/llms_still_cannot_plan_it_is_one_of_the_most/. Retrieved on December 15, 2024.

[33]: Reddit, (2024), “Why AGI can't be achieved with the LLM-based approach”, r/ArtificialInteligence, https://www.reddit.com/r/ArtificialInteligence/comments/1f8wnk9/why_agi_cant_be_achieved_with_the_llmbased/. Retrieved on December 15, 2024.

[34]: tobycrisford, (2024), “LLMs won't lead to AGI - Francois Chollet”, Effective Altruism Forum, https://forum.effectivealtruism.org/posts/um7tDLrjLWKG58vYa/llms-won-t-lead-to-agi-francois-chollet, June 11, 2024.

[35]: Hacker News, (2023), “The future of generative AI is niche, not generalized”, https://news.ycombinator.com/item?id=35730327. Retrieved on December 15, 2024.

[36]: Gonçalo Ribeiro, (2024), “Understanding The Limitations Of Generative AI”, Forbes, https://www.forbes.com/councils/forbestechcouncil/2024/05/09/understanding-the-limitations-of-generative-ai/, May 9, 2024.

[37]: Darko Matovski, (2024), “Causal AI: the revolution uncovering the 'why' of decision-making”, World Economic Forum, https://www.weforum.org/stories/2024/04/causal-ai-decision-making/, April 9, 2024.

[38]: NeuroCortex.AI, (2024), “Navigating the Landscape: Ethics, Governance, and Challenges in Generative AI”, https://medium.com/@theagipodcast/navigating-the-landscape-ethics-governance-and-challenges-in-generative-ai-bf8c8b8d5056, June 7, 2024.

[39]: Scott McLean, Gemma J. M. Read, Jason Thompson, Chris Baber, Neville A. Stanton, Paul M. Salmon, (2021), “The risks associated with Artificial General Intelligence: A systematic review”, https://www.researchgate.net/publication/353888549_The_risks_associated_with_Artificial_General_Intelligence_A_systematic_review, August 13, 2021.

[40]: Banghua Zhu, Norman Mu, Jiantao Jiao, David Wagner, (2024), “Generative AI Security: Challenges and Countermeasures”, arXiv:2402.12617v1.

[41]: Albert Meige, Zoe Huczok, Rick Eagar, Vincent Benyamin-Wood , Saloni Mehta, Leo Liu, Obvious, Luc Ferry, (2023), “Generative artificial intelligence: Toward a new civilization?”, Arthur D. Little, https://www.adlittle.com/en/insights/report/generative-artificial-intelligence-toward-new-civilization, October, 2023.

[42]: Catherine Dee, (2024), “The Ethics of Generative AI: Understanding the Principles and Risks”, Neo4j, https://neo4j.com/blog/genai-ethics/, September 18, 2023.

[43]: Lina Bariah and Merouane Debbah, (2024), “AI Embodiment Through 6G: Shaping the Future of AGI” - IEEE Wireless Communications Volume: 31, Issue: 5, October 2024.

[44]: What lies behind AGI: ethical concerns related to LLMs, accessed on December 15, 2024, https://montrealethics.ai/what-lies-behind-agi-ethical-concerns-related-to-llms/

[45]: The Ethics of Artificial General Intelligence (AGI): Navigating the Path to Human and Machine Coexistence. | by Gaurav Sharma | Medium, accessed on December 15, 2024, https://medium.com/@gaurav.sharma/the-ethics-of-artificial-general-intelligence-agi-navigating-the-path-to-human-and-machine-1ae571165f28

[46]: Giada Pistilli, (2022), “What lies behind AGI: ethical concerns related to LLMs”, Revue Ethique et Numérique, https://hal.science/hal-03607808.

[47]: Kyryl Sidak, (2024), “How will AGI Change Our Future”, Codefinity, https://codefinity.com/blog/How-will-AGI-Change-Our-Future, April, 2024.

[48]: Julia McCoy, “How to Prepare for AGI: Your Guide to the Future”, First Movers AI, https://firstmovers.ai/how-to-prepare-for-agi/. Retrieved on December 15, 2024.

[49]: Stephane H. Maes, (2024), “MultiAI Document Generation Optimized for task and context”, Internal Report, Factwise.ai, August 12, 2024.

[50]: Stephane H. Maes, (2024), ” MultiAI Document Generation Optimized for task and context at Intelligine.ai”, Intelligine.ai, October 1, 2024.

[51]: Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al., (2023), “A survey of large language models”, arXiv:2303.18223v15.

[52]: Yann LeCun, (2014), “"LLMs are an offramp on the path to AGI" […]”, Twitter, https://publish.twitter.com/?url=https://twitter.com/ylecun/status/1801018194121118103#, June 12, 2024.

[53]: Francois Chollet, (2024), “OpenAI has set back the progress towards AGI by 5-10 years because frontier research is no longer being published and LLMs are an offramp on the path to AGI”, Twitter, , https://publish.twitter.com/?url=https://twitter.com/ylecun/status/1801018194121118103#, June 12, 2024.

[54]: Thomas Macaulay, (2024), “Meta’s AI chief: LLMs will never reach human-level intelligence. Sorry, Elon — AGI won't arrive next year”, TNW, https://thenextweb.com/news/meta-yann-lecun-ai-behind-human-intelligence, April 10, 2024.

[55]: Stephane H. Maes, (2023), “Comments to MPAI-MMC V2 Draft – Request for Public Comments”, https://shmaes.wordpress.com/2023/09/16/comments-to-mpai-mmc-v2-draft-request-for-public-comments/, https://osf.io/hj6pw, September 16, 2023.

[56]: Stephane H. Maes, (2023), “Comments to MPAI-MMC V2 Draft – Request for Public Comments”, viXra:2310.0015v1, September 16, 2023.

[57]: Stephane H Maes, (2023), “A GUIDE TO Building Your Own Service Desk Virtual Agent and simulating human conversations”, https://www.researchgate.net/publication/370100099_Building_Your_Own_Service_Desk_Virtual_Agent_and_simulating_human_conversations_A_Guide_to_Building_Your_Own_Service_Desk_Virtual_Agent_and_Simulating_Human_Conversations, April 18, 2023.

[58]: Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, Denny Zhou, (2022-2023), “Chain-of-Thought Prompting Elicits Reasoning in Large Language Models”, arXiv:2201.11903v6.

[59]: Liam Sturgis, (2024),”Chain of Thoughts (COT) Prompting for LLMs: 7 Powerful Prompting Techniques to Get the Results You Want”, IBSN: 979-8884109520.

[60]: Lijie Hu, Liang Liu, Shu Yang, Xin Chen, Zhen Tan, Muhammad Asif Ali, Mengdi Li, Di Wang, (2024), “Understanding Reasoning in Chain-of-Thought from the Hopfieldian View”, arXiv:2410.03595v1.

[61]: Stephane H. Maes, (2023), “AI-ITSM: ChatGPT and IFS assyst”, IFS report.

[62]: Stephane H. Maes, (2022), “Smart ITSM, ESM, ITOM: adding AI, including Conversational AI, Beyond it, AIOPs and Beyond towards the Autonomous Enterprise”, IFS report.

[63]: Daniel Coffman, Liam D Comerford, Steven DeGennaro, Edward A Epstein, Ponani Gopalakrishnan, Stephane H Maes, David Nahamoo, (1998 to 2011), “CONVERSATIONAL COMPUTING VIA CONVERSATIONAL VIRTUAL MACHINE”, US patent US 8,082,153 B2.

[64]: Daniel Coffman, Liam D Comerford, Steven DeGennaro, Edward A Epstein, Ponani Gopalakrishnan, Stephane H Maes, David Nahamoo, (1998 to 2011), "CONVERSATIONAL COMPUTING VIA CONVERSATIONAL VIRTUAL MACHINE", US patent US 8,082,153 B2.

[65]: Stephane H. Maes, TLV Raman, (1997 to 2010), "Methods and systems for multi-modal browsing and implementation of a conversational markup language", US Patent 7,685,252.

[66]: Daniel Coffman, Liam D Comerford, Steven DeGennaro, Edward A Epstein, Ponani Gopalakrishnan, Stephane H Maes, David Nahamoo, (1998 to 2011), "CONVERSATIONAL COMPUTING VIA CONVERSATIONAL VIRTUAL MACHINE", US patent US Patent 7,137,126.

[67]: Stephane H. Maes, (2000), "Elements of conversational computing-A paradigm shift", International Conference on Spoken Language.

[68]: A Tiwari, RA Hosn, SH Maes, (2003), "Conversational multi-modal browser: an integrated multi-modal browser and dialog manager", 2003 Symposium on Applications and the Internet. Proceedings, 348-351.

[69]: J Gergic, J Kleindienst, S Maes, T Raman, J Sedivy, (2001-2003), "Systems and methods for providing conversational computing via javaserver pages and javabeans", US Patent App. 09/837,024.

[70]: Stephane H. Maes (1999), "Conversational biometrics", The European Conference on Speech Communication and Technology.

[71]: Stephane H. Maes, (2024), "From MultiAI to SingleAI honed to a specific domain, user/enterprise, and/or specialized for a specific type of document or media", Intelligine.ai, November 2024.

[72]: Fods12, (2023), "Intrinsic limitations of GPT-4 and other large language models, and why I'm not (very) worried about GPT-n", Effective Altruism Forum, https://forum.effectivealtruism.org/posts/6dphu3p8d5mQZEZzk/intrinsic-limitations-of-gpt-4-and-other-large-language, June 3, 2023.

[73]: Pablo Villalobos, Anson Ho, Jaime Sevilla, Tamay Besiroglu, Lennart Heim, Marius Hobbhahn, (2022 & 2024), "Will we run out of data? Limits of LLM scaling based on human-generated data", arXiv:2211.04325v2.

[74]: Zhyar Rzgar K Rostam, Sándor Szénási, Gábor Kertész, (2024), "Achieving Peak Performance for Large Language Models: A Systematic Review", arXiv:2409.04833v1.

[75] restack, (2024), "Trends In Large Language Models Performance", https://www.restack.io/p/large-language-models-answer-performance-trends-cat-ai, December 16, 2024. Retrieved on December 22, 2024.

[76] Educating Silicon, (2024), "How much LLM training data is there, in the limit?", https://www.educatingsilicon.com/2024/05/09/how-much-llm-training-data-is-there-in-the-limit/, March 9, 2024.

[77] Justborn, (2024), "Fine-Tuning Your Large Language Model for Peak Performance", https://justoborn.com/large-language-model/, May 2, 2024.

[78] Zhuoqiao Hong, et al., (2024), "Scale matters: Large language models with billions (rather than millions) of parameters better match neural representations of natural language", https://pmc.ncbi.nlm.nih.gov/articles/PMC11244877/.

[79]: Avril Aysha, (2023), "Improve your machine learning life cycle with synthetic data", https://mostly.ai/blog/machine-learning-life-cycle-with-synthetic-data, August 8, 2023.

[80]: Adam Zewe, (2022), "In machine learning, synthetic data can offer real performance improvements. Models trained on synthetic data can be more accurate than other models in some cases, which could eliminate some

privacy, copyright, and ethical concerns from using real data", https://news.mit.edu/2022/synthetic-data-ai-improvements-1103, November 2, 2023.

[81]: Cade Daniel, Chen Shen, Eric Liang and Richard Liaw, (2023), "How continuous batching enables 23x throughput in LLM inference while reducing p50 latency", https://www.anyscale.com/blog/continuous-batching-llm-inference, June 22, 2023.

[82]: Jiashuo Sun, Yi Luo, Yeyun Gong, Chen Lin, yelong shen, Jian Guo, Nan Duan, (2023 - 2024), "Enhancing Chain-of-Thoughts Prompting with Iterative Bootstrapping in Large Language Models", https://openreview.net/pdf?id=LIW88mwqgv.

[83]: Sandeep Kumar, (2023), "Problems with Synthetic Data", aitude, https://www.aitude.com/problems-with-synthetic-data/, March 1, 2023.

[84]: Michael Abramov, (2024), "Synthetic data definition: Pros and Cons", https://keymakr.com/blog/synthetic-data-definition-pros-and-cons/, October 16, 2024.

[85]: Wikipedia, "Causal AI", https://en.wikipedia.org/wiki/Causal_AI. Retrieved on December 18, 2024.

[86]: LeewayHertz, "From good to great: Enhancing your large language model's performance for desired outputs", https://www.leewayhertz.com/better-output-from-your-large-language-model/. Retrieved on December 15, 2024.

[87]: Prasad Deshmukh, (2023), "Synthetic Data and its impact on Machine Learning Models.", https://www.linkedin.com/pulse/synthetic-data-its-impact-machine-learning-models-prasad-deshmukh, February 21, 2023.

[88]: Shuang Hao, Wenfeng Han, Tao Jiang, Yiping Li, Haonan Wu, Chunlin Zhong, Zhangjun Zhou, He Tang, (2024), "Synthetic Data in AI: Challenges, Applications, and Ethical Implications", arXiv:2401.01629v1.

[89]: Ilia Shumailov, Zakhar Shumaylov, Yiren Zhao, Yarin Gal, Nicolas Papernot, Ross Anderson, (2023 & 2024), "AI models collapse when trained on recursively generated data", arXiv:2305.17493v3.

[90]: Ali Borji, (2014), "A Note on Shumailov et al. (2024): `AI Models Collapse When Trained on Recursively Generated Data'", arXiv:2410.12954v2.

[91]: Stephane H. Maes, (2024), "Fixing Reference Hallucinations of LLMs", https://doi.org/10.5281/zenodo.14543939, https://shmaes.wordpress.com/2024/11/29fixing-reference-hallucinations-of-llms/, November 29, 2024. (osf.io/u38w4/, viXra:2412.0149v1).

[92]: OpenAI et al., (2023 & 2024), "GPT-4 Technical Report", arXiv:2303.08774, V6 + from V1 through V6.

[93]: "LLM Document Extraction: How to use AI to get structured data from legacy documents", PondHouse Data, https://www.pondhouse-data.com/blog/document-extraction-with-llms. Retrieved on December 15, 2024.

[94]: Superuser, (2021), "How to auto-detect and export all the web links from a MS word file to a separated file?", StackExchange, https://superuser.com/questions/1671109/how-to-auto-detect-and-export-all-the-web-links-from-a-ms-word-file-to-a-separat. Retrieved on December 15, 2024.

[95]: Michael Timothy Bennett, Elija Perrier, (2024), "An AI system has reached human level on a test for 'general intelligence'. Here's what that means", The Conversation, https://theconversation.com/an-ai-system-has-reached-human-level-on-a-test-for-general-intelligence-heres-what-that-means-246529, December 24, 2024.

[96]: François Chollet, (2024), "OpenAI o3 Breakthrough High Score on ARC-AGI-Pub", ArcPrize, https://arcprize.org/blog/oai-o3-pub-breakthrough, December 20, 2024.

[97]: Freedom Preetham, (2024), "Part 1 — Why LLMs Will Never Lead to AGI. The Mathematical and Biological Barriers", Medium, https://medium.com/autonomous-agents/why-llms-will-never-lead-to-agi-aa7bcff9805d, September 26, 2024.

[98]: Chia Xin Liang, Pu Tian, Caitlyn Heqi Yin, Yao Yua, Wei An-Hou, Li Ming, Tianyang Wang, Ziqian Bi, Ming Liu, (2024), "A Comprehensive Survey and Guide to Multimodal Large Language Models in Vision-Language Tasks", arXiv:2411.06284v1.

[99]: Shukang Yin, Chaoyou Fu, Sirui Zhao, Ke Li, Xing Sun, Tong Xu, Enhong Chen, (2023 & 2024), "A survey on multimodal large language models", arXiv:2306.13549v4.

[100]: Stephane Maes, (2020-24), "Web Site Tracking all Publications around the Multi-fold universe", Navigation page listing all papers, https://shmaesphysics.wordpress.com/shmaes-physics-site-navigation/.

[101]: Stephane H. Maes, (2020-2022) "Quantum Gravity Emergence from Entanglement in a Multi-Fold Universe", HIJ, Vol 2, No 4, pp 136-219, Dec 2022, https://doi.org/10.55672/hij2022pp136-219, https://shmaesphysics.wordpress.com/2020/06/09/paper-published-as-preprint-quantum-gravity-emergence-from-entanglement-in-a-multi-fold-universe/, https://shmaesphysics.wordpress.com/2022/11/09/quantum-gravity-emergence-from-entanglement-in-a-multi-fold-universe-2/, and viXra:2006.0088, (June 9, 2020). Errata/improvements/latest updates at https://zenodo.org/doi/10.5281/zenodo.7792911.

[102]: Stephane Maes, (2020-24), "Web Site Tracking all Publications around the Multi-fold universe", Navigation page listing all papers, https://shmaesphysics.wordpress.com/shmaes-physics-site-navigation/.

[103]: Stephane H Maes, (2021), "The Multi-fold Theory: A synopsis", viXra:2112.0144v1, https://shmaesphysics.wordpress.com/2021/12/24/the-multi-fold-theory-a-synopsis-so-far-v2-end-of-2021/, December 24, 2021. Note that additional links will always be available at https://shmaesphysics.wordpress.com/2021/05/03/the-multi-fold-theory-a-synopsis-so-far/ to track the latest and interim versions of the synopsis, as they may be published under different tittle or URL/publication numbers.

[104]: Stephane H Maes, (2022), "Understanding the Multi-fold theory principles and the SM_G", osf.io/xc74t, https://shmaesphysics.wordpress.com/2022/03/11/understanding-the-multi-fold-theory-principles-and-the-sm_g/, March 11, 2022. Also as Stephane H Maes, (2022), "A tutorial on the Multi-fold theory principles and the SM_G", viXra:2303.0154v1, https://shmaesphysics.wordpress.com/blog-2/a-tutorial-on-the-multi-fold-theory-principles-and-the-sm_g/, March11, 2022.

[105]: Stephane H. Maes, (2022), "What is the Multi-fold Theory? Its Main Characteristics in a Few Words", vixra:2207.0172v1, https://shmaesphysics.wordpress.com/2022/07/28/what-is-the-multi-fold-theory-its-main-characteristics-in-a-few-words/, July 28, 2022.

[106]: Stephane H Maes (2023), "The Multi-fold Theory – Draft Raw Compendium of Research Papers (till August, 2023)", https://doi.org/10.5281/zenodo.8242021, https://shmaesphysics.wordpress.com/2023/08/12/the-multi-fold-theory-draft-raw-compendium-of-research-papers-till-august-2023/, August 12, 2023, (https://osf.io/swqmb).

[107]: Stephane H. Maes, (2020-2023), "Quantum Gravity Emergence from Entanglement in a Multi-Fold Universe", V3, https://zenodo.org/doi/10.5281/zenodo.7792911, October 29, 2023.

[108]: Stephane H. Maes, (2024), "June 2024 Status of the Multi-fold Theory", https://doi.org/10.5281/zenodo.13345964, https://shmaesphysics.wordpress.com/2024/08/19/june-2024-status-of-the-multi-fold-theory/, June 22, 2024.

[109]: Asif Razzaq, (2024), "Microsoft Open Sourced MarkItDown: An AI Tool to Convert All Files into Markdown for Seamless Integration and Analysis", https://www.marktechpost.com/2024/12/18/microsoft-open-sourced-

markitdown-an-ai-tool-to-convert-all-files-into-markdown-for-seamless-integration-and-analysis/, December 18, 2024.

[110]: Fahd Mirza, (32024), “Microsoft MarkItDown - Convert Files and Office Documents to Markdown - Install Locally”, YouTube, https://www.youtube.com/watch?v=v65Oyddfxeg. Retrieved on December 22, 2024.

[111]: Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, Denny Zhou, (2022-2023), “Chain-of-Thought Prompting Elicits Reasoning in Large Language Models”, arXiv:2201.11903v6.

[112]: Liam Sturgis, (2024),”Chain of Thoughts (COT) Prompting for LLMs: 7 Powerful Prompting Techniques to Get the Results You Want”, IBSN: 979-8884109520.

[113]: Lijie Hu, Liang Liu, Shu Yang, Xin Chen, Zhen Tan, Muhammad Asif Ali, Mengdi Li, Di Wang, (2024), “Understanding Reasoning in Chain-of-Thought from the Hopfieldian View”, arXiv:2410.03595v1.

[114]: Maxwell Zeff, (2024), “OpenAI trained o1 and o3 to ‘think’ about its safety policy”, https://techcrunch.com/2024/12/22/openai-trained-o1-and-o3-to-think-about-its-safety-policy/, December 22, 2024.

[115]: Freedom Preetham, (2024), “Part 4 — A Mathematical Framework for Fluid Intelligence”, https://medium.com/autonomous-agents/part-4-mathematical-framework-for-fluid-intelligence-7a2fa36cd004, October 27, 2024.

[116]: Wikipedia, “Fluid and crystallized intelligence”, https://en.wikipedia.org/wiki/Fluid_and_crystallized_intelligence. Retrieved on December 22, 2024.

[117]: Stephane H. Maes, (2022), “CO2 and CH4 absorption powered by nuclear fusion, via fission, is the only way to manage climate change and the Planet’s trigger points”, viXra:2211.0154v1, https://shmaes.wordpress.com/2022/04/09/co2-and-ch4-absorption-powered-fission-is-the-only-way-to-manage-climate-change-and-the-planets-trigger-points/, April 9, 2022, https://osf.io/5ymds.

[118]: Stephane H. Maes, (2024), “The environmental cost of GenAI is out of control. It triples emissions of data centers!”, https://shmaes.wordpress.com/2022/04/09/co2-and-ch4-absorption-powered-fission-is-the-only-way-to-manage-climate-change-and-the-planets-trigger-points/#comment-641, September 10, 2024, and other comments on [117].

[119]: Wikipedia, “Knowledge distillation”, https://en.wikipedia.org/wiki/Knowledge_distillation. Retrieved on November 28, 2024.

[120]: Zhen Huang, Haoyang Zou, Xuefeng Li, Yixiu Liu, Yuxiang Zheng, Ethan Chern, Shijie Xia, Yiwei Qin, Weizhe Yuan, Pengfei Liu, (2024), “O1 Replication Journey — Part 2: Surpassing O1-preview through Simple Distillation, Big Progress or Bitter Lesson?”, arXiv:2411.16489v1.

[121]: Stephane H Maes, TV Raman, (2000), “Multi-modal interaction in the age of information appliances”, 2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No. 00TH8532), July 30, 2000.

[122]: Stephane H Maes, TV Raman, (2000), “Multi-modal Web IBM Position”, W3C/WAP Workshop, IBM Human Language Technologies, September, 2000.

[123]: Stephane H Maes, (2001), “Multi-modal Architecture Recommendation and Multi-device Browsing”, Proceedings of the 6th National Conference on Human-Computer Speech Communication (NCMMSC6), 2001, 2001.

[124]: Stephane H Maes, Chalapathy Venkata Neti, (2005), “System and method for multi-modal focus detection, referential ambiguity resolution and mood classification using multi-modal input”, US Patent 6964023, November 8, 2005.

[125]: David Boloker, Rafah A Hosn, Photina Jaeyun Jang, Jan Kleindienst, Tomas Macek, Stephane H Maes, Thiruvilwamalai V Raman, Ladislav Seredi, (2006), “Systems and methods for implementing modular DOM (Document Object Model)-based multi-modal browsers”, US Patent 7028306, April 11, 2006.

[126]: John M Lucassen, Stephane H Maes, (2006), “MVC (model-view-controller) based multi-modal authoring tool and development environment”, US Patent 6996800, February 7, 2006.

[127]: Jennifer A Healey, Rafah A Hosn, Stephane Herman Maes, Thiruvilvama Lai V Raman, Alpana Tiwari, (2009), “Methods and systems for authoring of mixed-initiative multi-modal interactions and related browsing mechanisms”, US patent 7546382.

[128]: Stephane H. Maes, (2013), “Platform and service for management and multi-channel delivery of multi-types of contents”, US Patent 8365306.

[129]: Peters, J., Janzing, D., & Schölkopf, B., (2017), :Elements of Causal Inference: Foundations and Learning Algorithms”, The MIT Press.

[130]: Bareinboim, E., Correa, J. D., Ibeling, D., Icard, T., (2022), “On the Causal Foundations of AI. Nature Machine Intelligence”, 4(7), 588-593.

[131]: Pearl, J., (2009), “Causality: Models, Reasoning, and Inference”, (2nd ed.). Cambridge University Press.

[132]: Ananya Gairola, (2024), “Microsoft Invested Nearly $14 Billion In OpenAI But Now Its Reducing Its Dependence On The ChatGPT-Parent: Report”, Benzinga, https://www.benzinga.com/24/12/42664762/after-investing-nearly-14-billion-in-openai-microsoft-reportedly-moves-to-reduce-its-dependence-on-chatgpt-parent, December 23, 2024.

[133]: Stephane Herman Maes, Karan Singh Chhina, Guillaume Dubuc, (2021), “Natural language translation-based orchestration workflow generation”, US Patent 11120217.