The Technology of Thinking/Doing: The Autonomous Agent Acts

Sarah Lilly Eaton, Ph.D. and Rajeev Bhargava, Founder
Decision-Zone, Inc.

With gratitude to the contributions of Dr. John Kenney
for his work in specifying the *Rapide* language and ongoing involvement with Decision-Zone

January 07, 2024

παράγει γὰρ τὸ σχῆμα τοῦ κόσμου τούτου.
For the present schema of this world is passing away.

Abstract

This paper explores the advancement of Artificial General Intelligence (AGI) from a philosophical and technological perspective. The pursuit of AGI is fundamentally about finding a scalable and effective way to model and replicate human-like intelligence. The limitations of current data-driven approaches are considered, emphasizing the limitations of current AI models. It emphasizes the importance of continuous interaction with the environment, the integration of causality at a foundational level and the capacity for autonomous real-time decision-making. Central to the paper's argument is the concept of "thinking now occurring," a philosophical orientation rooted in the work of David G. Leahy and drawing upon his critique of American pragmatism to collapse traditional dualisms in cognition. Drawing upon Leahy's articulation of world pragmatism it advocates for a process-centric approach to AGI. Further, it emphasizes the importance of temporality in human experience and argues for the need to integrate this into AGI development. The key claim is that *for a thinking machine to be able to operate in the now, just as the human thinker does, there must be an intelligibility that is not incompatible with immediacy*. The core of the paper is dedicated to detailing the technological framework of DADA X, including its event-driven architecture, the *Rapide* language, and the significance of activity models in enabling autonomous decision-making. This platform is presented as a solution to the challenges of current AI approaches, enabling more sophisticated, real-time interactions with the environment. The conclusion points towards future research directions, including a fuller treatment of ethical modeling, decentralization and the development of on-device AGI capabilities.

1 Introduction

We will be considering artificial general intelligence (AGI) from a philosophical perspective, understanding that "philosophy itself consists of conceptual investigations whose essential nature is neither empirical nor logico-mathematical." [1] The purpose of this high level conceptual consideration is to articulate the philosophical approach that formed the basis of development of the DADA X platform, and to justify our stance that the path forward to AGI requires core technological capabilities that currently only exist with DADA X.

In evaluating AGI systems, Google DeepMind has recommended the principle: Focus on potential, not deployment. They explain that "defining AGI in terms of reaching a certain level of labor substitution would require real-world deployment, whereas defining AGI in terms of being capable of substituting for labor would focus on potential." [2] Our discussion will outline the potential of our platform for creating and powering the kinds of applications that are needed for AGI to become a reality.

We align with a functionalist view of intelligence as the ability to perform cognitive tasks, emphasizing problem-solving, self-correction, adaptation, and decision-making. Intelligence is not knowledge in a vacuum but is conceived holistically as perceiving-thinking-acting. Although a working definition of AGI is sufficient for limiting our reference, the purpose in this paper goes further in an attempt to explore at a conceptual level the core mechanism required to progress toward the goal.

1.1 The Stakes

The race towards Artificial General Intelligence is the contemporary equivalent of the Manhattan Project, and the stakes are exceptionally high. According to Prof. Zhu Song-Chun, who leads the Chinese efforts to secure global leadership in AGI: "Artificial general intelligence (AGI) is the strategic high ground of international scientific and technological competition in the next ten to twenty years, and its influence is equivalent to the 'atomic bomb' in the field of information technology." [3]

The urgency of the situation cannot be overstated. Whoever gains the upper hand in AGI will hold a potentially world-dominating technology. With China's governmental support and focus, other countries are at risk of falling behind unless they make AGI a national priority.

1.2 Definitions of AGI

What is required in the definition. Because AGI has been one of the most disputed concepts in technology, it is useful to examine some current definitions offered by thought-leaders and corporations in the realm of AGI. These definitions are all focused on what a future AGI system should be able to *do*. The question of how these abilities are to be enabled is bound up in various assumptions which will be discussed later. Google DeepMind has proposed, among other criteria, that any definition of AGI should "focus on what an AGI can accomplish, not on the mechanism by which it accomplishes tasks." [4] This criteria is intended to avoid the implication that an AGI must *understand*, be *conscious*, or even *sentient* since those are not measurable. But let us not be shy to admit that all the definitions of AGI, insofar as they describe "cognitive abilities" are describing a thinking machine.

In some contrast to Google's dictum, we will not avoid but welcome discussion of the means by which the capabilities can be implemented in computing, as these are not settled matters. In the present context, our mechanism for enabling the capabilities is precisely at issue and one which does not require any assumptions about the machine being conscious or sentient. However, it will eventually lead into some discussion of what is meant by "understanding". Words like "learning" and "reasoning" are often used in definitions of AGI cognitive abilities and "understanding" is no more or less problematic in our view than these notions. In any case, it is no less problematic than "thinking" itself.

Exemplary definitions. It is important to keep in mind that since AGI is currently just a future possibility, the question is not 'what can it do?', but 'what do we want it to do?' The following definitions are thus reflections of the desires of the human beings who hope to create them.

DeepMind founder Shane Legg, in an October 2023 interview stated that "When I say AGI, I mean a machine that can do the sorts of cognitive things that people can typically do, possibly more." [5]. This understanding is more thoroughly explored in a recent paper from DeepMind which attempts to provide a framework for classifying the capabilities and behaviors of AGIs. "Whether to require robotic embodiment (Roy et al., 2021) as a criterion for AGI is a matter of some debate. Most definitions focus on cognitive tasks, by which we mean non-physical tasks." [6]

OpenAI's charter is more focused, defining AGI as "highly autonomous systems that outperform humans at most economically valuable work". [7] We appreciate OpenAI's inclusion of the concept of autonomy within the definition and will focus more on this aspect in our discussion.

Prominent AGI researcher Ben Goertzel, CEO of SingularityNet, is often credited for first using the phrase "artificial general intelligence" in a book published almost 20 years ago. He holds that intelligence may be at least partially conceived as "the ability to achieve complex goals in complex environments" where complexity is the "possession of a rich variety of patterns." He includes the capacity for reflection in his notion of intelligence and sees an intelligent system "as a dynamical system that recognizes patterns in its environment and itself, as part of its quest to achieve complex goals." [8] Rather than merely human-level intelligence, he goes further and sees AGI as surpassing human capabilities by making inferences outside its initial training data. He foresees an AGI that can "reflect on its own structure, to reason about its own operations and perceptions in a combined way and to intelligently guide self-modifications." [9] In some ways, what Goertzel describes under the umbrella of AGI may be described by others as ASI (artificial super intelligence).

Joscha Bach defines nine capability dimensions or vectors of intelligence: perception, learning, representation, reasoning, knowledge, language, autonomy, collaboration and embodiment. AGI is then the "system that reaches human level or beyond at all capability dimensions." [10]

Gary Marcus, a cognitive scientist and leading voice in the field of AI who is skeptical of current approaches, has stated that AGI would require systems that: essentially never hallucinate, reliably reason over abstractions, can form long term plans, understand causality, reliably maintain models of the world, and reliably handle outliers. ("We currently have none of that," he tweeted). [11]

Another definition, offered by Peter Voss in his article "What is AGI", encapsulates the grand vision as follows: "a single system that can learn incrementally, reason abstractly, and act effectively over a wide range of domains — just like humans can." Voss explains that AGI is specified in terms of human abilities because the goal is to free us from most work. "The core problem is building the brain, the intelligence engine." [12] This is the central task, beyond which the requirements for sensory connection to the environment can then be gradually expanded to allow physical operations and tool use. By contrast, Meta's Yann LeCun has argued that artificial intelligence needs to learn how the world works from multi-modal sensory inputs, making the sensory connection integral from the outset in order to build up an implicit world model.

1.3 Our Working Definition of AGI

It is the pragmatic goal that must be kept in mind as we explore the path forward to AGI using the decentralized autonomous decisioning agent platform. An exhaustive enumeration of tasks is not required; as Google DeepMind states: “It is impossible to enumerate the full set of tasks achievable by a sufficiently general intelligence.” [13]

AGI is a distributed system capable, minimally, of performing human-level cognitive tasks autonomously and, maximally, of executing a broader range of tasks within the environment when ‘embodied’ in robots or other devices. AGIs as avatars would involve the addition of a model discovery/generation algorithm and other perceptual capabilities such as vision and language that would broaden the ability to interact by means of an explicitly modeled world of activities and processes. Anything falling short of agency with autonomous correction and control capabilities will not reach human level cognition. Dynamic, real-time decision-making is central to achieving the level of performance expected from AGIs.

1.4 Challenges for Achieving General Intelligence

Is Increasing Compute the Answer? We cannot make the assumption that more massive models and greater computational power would lead to Artificial General Intelligence. Although increased compute was indeed crucial for the development and operation of applications like GPT4, it is now widely suggested that it may have reached its limits in advancing AI capabilities. Despite billions invested in other LLMs, none have been able to improve upon GPT4 by any order of magnitude thus far.

Are Large Language Models a path to AGI? It might seem obvious that text prediction alone is not a sufficient basis for Artificial General Intelligence as it has been defined above. Nevertheless, some have been excited by the so-called “sparks” of AGI glimpsed when OpenAI rolled out it’s ChatGPT to the public more than a year ago. [14] The novel and impressive abilities of LLMs have captivated us even to the point of fear, leading thousands to sign an open letter asking for a pause in research (which notably did not happen). Unfortunately, the success of those models in performing the task of predicting the next word in a given context threatens to blind us to other approaches.

In the past year, the limitations of LLMs have been more thoroughly adumbrated. Most crucially the focus has been on the fact that the intrinsic nature of the models results in hallucination/confabulation— leading users to question the accuracy and reliability of the outputs. A year ago, optimistic researchers believed the hallucination problems might be solved in a few months’ time through scaling alone, a hope borne perhaps from the idea that scaling is easier than thinking. Despite the fact that finding ways to mitigate hallucination has become an obsession, the issue is unresolved. Ilya Sutskever, Chief Scientist of OpenAI, was asked in an interview in March 2023 to speculate on the chief reason that the GPT approach may turn out not to succeed. His answer: unreliability. [15]

Many researchers say that the problem is inherent to their design— that hallucination is the fundamental feature, not the bug. This may be appropriate for certain creative projects but the wrong kind of technology for outputs that must be truly trustworthy. Verifying and checking for hallucination is in itself a burdensome task, which leads us to doubt the overall productivity gain of using LLMs. Transparency initiatives which aim to make AI's 'black-boxes' more understandable do not inherently improve the reliability of generative AIs.

Not only are the financial and reputational risks to business high, but the social consequences run deep. According to Dr. Mark van Rijmenam, generative AI and deepfakes are shaping a "post-truth" and "post-trust" world. The erosion is occurring now. Clearly there is a need for a better solution that improves the accuracy and reliability of language models as "distinguishing fact from fiction becomes an intricate task, posing a significant challenge to individuals and institutions and having a major impact on democracies worldwide." [16]

Coherence is also a problem. Joscha Bach has simply characterized the difference between LLMs and minds in this way: LLMs are able to complete sequences (predict next tokens) but minds have understanding which means mapping to a universal model. Patterns without the model are at risk of incoherence. Not only are LLM patterns not mapped to a universal model, they are not designed to link the token patterns to corresponding documents in the training corpus from which they were drawn. Furthermore, he reminds us that "the game is not about intelligence it's about agency, it's about the ability to control the world and intelligence is instrumental to this… so it's a very different way of looking at things than the current language models which have basically no agency." [17]

The stumbling block in the hypothetical path to AGI from LLMs is that in their basic form they are tools, processing user inputs but without proactive capability. There may be many useful tools in the toolbox, but AGI must wield them itself. OpenAI's approach to AGI was to build the language model first before tackling the matter of agency, which is arguably much more difficult. To solve complex, real world problems, and to achieve human level general intelligence requires so much more: activity models, world models, causal understanding and autonomy. Can we expect these capabilities to simply emerge?

1.5 Emergence: Miracle or Mirage?

Emergence is the phenomenon of new properties manifesting in a system as its complexity increases. Goertzel's patternist view is that the intelligence of a system includes "emergent patterns" as well, arising out of interactions with other systems.

With respect to LLMs, Herbert L. Roitblat, Principal Data Scientist at Mimecast, makes the point: "It would be extraordinary if advanced cognitive processes were to emerge from a word-guessing model. Extraordinary claims require support by extraordinary evidence, that evidence has not yet been forthcoming." [18]. In fact, evidence points in the other direction. The best paper award at the recent NeurIPs 2023 conference was given to a paper which argued that the so-called emergent capabilities of LLMs do not appear as a result of scale, but are simply attributable to the researcher's choice of metric. [19]

The idea of emergence is the technological analogue of an evolutionary explanation. Whether or not human intelligence arrived on the scene over millions of years of evolution or through intelligent design, the fact is that human intelligence is capable of designing intelligence without simply assembling components, crossing our fingers for good luck, and hoping that something intelligent emerges or evolves out of it. There's no reason to wait for this. We can design intelligence into our machines, not rely on emergent properties to achieve AGI. Whether machines that exhibit general cognitive abilities at human-level are or are not conscious has not been a primary consideration here. No strong claim is being made as to whether consciousness itself is to be expected as an emergent epiphenomena of a sufficiently complex AGI. It is quite plausible that the emergence of advanced cognitive processes is less probable in narrowly designed AI models, compared to the potential for consciousness to arise from an autonomous machine intelligence that is both model-based and event-driven.

2 Philosophical Considerations

Orientation. The American pragmatist philosophical movement which began in the 1870s and continued into the early part of the 20th century with thinkers such as William James, Charles S. Peirce and John Dewey, epitomized the kind of thinking which drove the innovation, progress and successes of the last century, leading to a world of increasingly advanced automation, expert systems, telecommunication and information processing technology, computing, etc.. In particular, the relevance of Peirce's philosophy to AI is evidenced by the substantial theoretical literature devoted to his semiotics and theory of representation "in the context of a critical completion of classical AI and/or classical computational theory of mind." [20]

The particular approach taken here has been articulated as the "Thinking Now Occurring" in distinction from the historical American pragmatist movement. As Dr. Steven Hoath explained: "This new thought termed by Leahy the thinking now occurring is a natural if unexpected outgrowth of American pragmatism extending the thought of Peirce, James, and Dewey from a merely relative objectivity to an absolute objectivity of consciousness." [21]

In the preface to his magnum opus, *Foundation: Matter the Body Itself*, David G. Leahy begins by stating that the beginning of the new world order is the end of modernity. New thinking is required, adequate to this new beginning. Yet philosophical treatments of artificial intelligence remain stuck in modernist categories, frameworks and paradigms. According to Leahy: "the new beginning is categorical… the categories and, indeed, the very structures of modern philosophical, theological, and scientific self-consciousness are essentially inadequate to the new beginning, and, further… the most fundamental structure, the very notion of *self*—in any but a purely formal sense—is completely and essentially dysfunctional in the light of the beginning of this new world." [22] Many critics of the hype surrounding current AI bemoan the "anthropomorphizing" of these tools, and rightly so. We must take care when focusing on the requirements for AGI that we project accurately what is and is not necessary for a machine to be able to think, based on our understanding of the nature of thinking unshackled by the weight of self-consciousness that is yet to be overcome.

To achieve AGI, a shift is needed towards the ultimate pragmatic view in which cognition is understood as a continuous flow of interactions with the environment, collapsing the oppositions: perception vs. thought, thought vs. action, and even mind vs. object. Not only do we find the pragmatic undermining of Cartesian dualism in Peirce, but also what Pierre Steiner calls the "post-cognitivist" notions of extended and embedded cognition. [23]

We are now on the threshold of an uncharted era which invites us to think anew and begin again. It is world pragmatism, the thinking first elucidated by Leahy, which inspires our approach to defining and explaining the 'thinking machine.' For world pragmatism, thinking's whole purpose is to construct the world. How can thought proceed in an unfettered way with the complete responsibility of creating the world? At this crucial juncture, by forcing, we risk creating tools and systems which ingest the data of humanity's subjective biases, flaws, and contentions only to magnify them exponentially.

Philosophical approaches to the question of intelligence have been attempts to understand what it is about the human "I" that underpins all these cognitive activities (*Ego cogito*). Our inquiry diverges significantly from the approach adopted by those who hope to uncover answers within the realm of neuroscience, a path distinct from the one we intend to elaborate here. Rather than focusing on the the neural architecture of human intelligence in the material structure of the brain, or even substantializing "Mind" in a dualist/Cartesian sense, our phenomenological and pragmatist account holds that "The I is absolute act." To model the "I" is to model activity, not a material structure. As Leahy puts it: "When 'I' think, this is not the work of a self. It is simply thinking that is occurring, specific and specifying, located in the now." [24]

2.1 Time

Conceptual discussions of AGI generally fail to do justice to a pivotal element intrinsic to human experience: temporality. Although I can check the time on my watch and say that at this moment it is 6:08pm, from the phenomenology of my experience it is always "now" for me. And in that now, I experience "just exactly this-ness."

As Leahy states: "Now is the quantum absolute." According to scientific research, which sought to answer the question: "how long is now?" humans experience a now of roughly 3 seconds (which seems quite long!) For a machine, this now is more difficult to characterize, depending upon what is being considered. Phenomenologically, one might argue that there neither is nor can be an 'experienced now' or 'what it is like to be now' for a machine. Another way of considering it would be to simply look at the processing speed of the hardware, how many fractions of a second does it take for the machine to execute. Rather than trying to exact upon it through a measurement of duration, consider that 'now' is the force of particularity, or as Leahy says, the fact of identity. He writes: "Time is not now except it be what now itself is, viz., the measure of the absolute contact of existence with thought." [25]. We extend his insight to claim that the thinking machine's measure of contact with existence is it's 'now.'

By contrast, the current paradigms in weak or narrow AI predominantly rely on static data, schemas and retrospective analyses. Such approaches engage with data that is inherently 'dead' — stored snapshots and traces of past events, devoid of the dynamic, flowing nature of real-time experience. The first problem is that historical data cannot fully explain the present. The second problem is that data has challenges in accurately representing and reconstructing the past. Predictive statistical models, anchored in sets of historical data are used to project possible future states but all are effectively representing what is no longer now and what is not yet now. How can we trust a prediction when the system cannot accurately determine what is happening in the present moment? Furthermore, can we trust a prediction when it is based on a record of the past that is filled with errors and noise? This disconnection, not only from the fluidity of time and but even from the discreteness and discontinuity of the moment of the now, results in a significant gap between human cognition and any cognition which would be imputed to today's artificial systems. Stated bluntly, current technology is not in the now to the extent that its “contact with reality” is mediated. It cannot handle a real-time, real-world environment. The orientation is based on objects, not activities, and simulation rather than true replication. Focus on the brute force of scaling with ever increasing memory, hits (and cannot hope to surmount) the wall blocking entry into contact with existence which is absolutely now. To do so, the system would need to be able to arrive at meaning in the discrete and discontinuous, not in post-facto analysis but in 1) the synthesis of the fact of the single event in its particularity and 2) the ability to determine it as correct or anomalous immediately. This perspicuous availability of the discrete thought/event to the system at the input side is conceivable with an extension of the philosophy of information theory.

2.2 Intelligibility Considered

There is a long history of assumptions that feed into the situation we find ourselves in with respect to how some believe intelligence could be achieved by our current systems. It is not a simple matter to reconstruct how we have arrived at where we are now, in terms of the total trajectory, so we take the shortcut, instead, by way of American pragmatism. It is instructive to return to the insights of Peirce to highlight the distinctions with the thinking now occurring. This will illuminate a view of intelligence, perception, cognition and their relationship to a new way of thinking about how AGI will be instantiated.

In a sense, we need not ask, primarily, ‘what is intelligence?’ but ‘what makes for intelligibility?’ What sort of interaction between the system and the inputs allows for intelligibility in the now (a synthesis of the fact) rather than mediated intelligibility (post-fact analysis)?

For Peirce, a thought, taken as a discrete thing, is without meaning. This is an important point which we would like to make more controversial as it seems that this notion has become rather ubiquitous. In the pragmatic view, only the continuum of the series arrives at meaning; meaning *is* only in the context of continuity. Neither meaning nor cognition arise *in the now*, so to speak. In addition, the development of thought is viewed in an evolutionary way (whether that evolution is seen at the personal level or the cosmological level). Semiotically, a sign and its interpretation constitute yet another sign, and this process is endless.

Central to Peirce’s insights are the categories of Firstness, Secondness and Thirdness and these triads are elaborated in different ways in various contexts. He writes: “The first is that whose being is simply in itself, not referring to anything nor lying behind anything. The second is that which is what it is by force of something to which it is second. The third is that which is what it is owing to things between which it mediates and which it brings into relation to each other." [26]

Firstness is the immediate, sheer givenness, freedom, sometimes described as a fullness of feeling without parts, a quality, pre-reflexive. In its givenness and immediacy, it is meaningless. Secondness is the force of actuality or existence, the hard fact of otherness. Thirdness is by nature relative. "A Third is something which brings a First into relation to a Second." [27] Leahy describes Peirce's Firstness as "the 'orience', being, or freedom, irreducible to self-consciousness," Secondness as "pure otherness, or existence as force or resistance" and Thirdness as "the middle term, relation, or essence." [28] The primacy of abductive inference in Peirce is related to the notion of thirdness.

The essence of American thought (epitomized by Peirce) is in Leahy's account characterized as absolute ideology (the pure relativity of the ideal consciousness with logic rooted in the social principle), the three categories of which are 'freedom', 'force' and 'mediation.' According to Peirce, logic requires that our interests should not be selfishly limited but extend to community of all beings with whom we can come into intellectual relation, even beyond "this geological epoch", even to the point of heroic self-sacrifice. For Peirce, the object of a person's love should ultimately be the community, but cannot be any particular existing community. The ideal quest is a thrust towards an ever-receding goal, an end never arrived at. Leahy calls this the "belief in the infinite self-postponement of technological consciousness: the belief in the technology of self-postponement, at once the belief in the self-postponement of technology." [29]

But according to the thinking now occurring, "the new world order categorically subsumes American thought." World-consciousness and world pragmatism begins as "the absolutely technological displacement of ideology: absolute technology absolutely displacing the belief in technology." To the extent that either "effective altruism" or "effective accelerationism" would subsume the foundation of social order in a self-sacrificial identification of the individual with a future, "unlimited community" they are both what Leahy calls "a blind seeing: the monad related to the other *ideally*." [30] Instead we should acknowledge that in the new order, new humanity, new consciousness and new thinking in which technology absolutely displaces ideology, the ideal of the "modern" consciousness, essentially past, is reduced to a new form: "the foundation of society is the immediately existing other." [31]

To see the implications of this at the level of the thinking machine we need to introduce a new concept to the scheme. For Leahy, Peirce's categories serve as the point of departure for a new category: Fourthness. In the thinking now occurring, Fourthness is a New Firstness qua foundation. It is described as follows:

> If Peirce distinguished Firstness, Secondness, and Thirdness, then add to these a Fourthness, an Absolute Now, Fact or Identity, in effect, an Absolutely Unconditioned Break in the Continuum, in which the three Peircean categories are themselves the elements not of "one undivided feeling without parts," but of "one undivided feeling" not without parts, no less undivided— the elements of the category of absolute quality—the object absolutely and immediately cognized as the now actual particular. [32]

For a thinking machine to be able to operate in the now, just as the human thinker does, there must be an intelligibility that is not incompatible with immediacy. At the lowest level of granularity an event in the stream of events must be identifiable in the 'now' of the machine, and cognized as the actual particular that it is. In the thinking now occurring, the availability of the discrete thought is immediate, synthesized, intelligible. Rather than describing the full range of activities that intelligence enables, it must begin to be thought in terms of grasping a single event in its fullest complexity to the point of being able to discern the correctness or the anomalous nature of that input. This shifts the focus from introspection and analysis to immediate experience, emphasizing the role of responsiveness in intelligent behavior. The identifiable pattern of a complex event is a whole with distinguishable parts, related to a larger model that allows for, as Leahy would say, a mediately immediate intelligibility when the pattern is instantiated particularly. Mediacy and Immediacy are an opposition that vanishes in the fact of identity, the Now.

2.3 The Idea of the World and the "World Model"

Some insist that the apparent successes of LLMs indicate that somewhere in the model, language itself has been enough to give rise to a world model within. But others have attributed the limitations of LLMs to their lack of a comprehensive world model. Yann LeCun has argued that sensory grounding is needed for meaning and understanding. His proposal is that when the model becomes multimodal, able to deal concurrently not only with text data but also images, video, audio and other inputs, a complex world model will emerge as a sort of technological analogue of common sense. Google's Gemini is the latest offering of a multimodal model.

We agree that AGI will indeed have a "world model," but the nature of such a model and how it is created is at issue. What, exactly, is being modeled? Is to model a world the equivalent of modeling a data generating process? The world is a field of action for engagement. And world order is subject to essential transformation. To reject this insight is to fall back on habit, tradition, and resist change and innovation. Is a model containing different types of data, integrated by algorithms, capable of modeling the dynamism of a world in which change is the fundamental constant? While an AGI must have various perceptual modalities, we contend that there are inherent limitations in current AI technology, in terms of both processing power and algorithmic sophistication needed for accurately modeling a dynamically changing environment such that human operating logic can be replicated in real world scenarios.

If AGI is described as goal-directed action, the idea of a world model as a process or activity model, based on causal and contextual event patterns is a more useful, and indeed, wholly pragmatic conception. The world model that is simulated then becomes more than a background against which the system operates, more than a virtual container of rules for the mechanical mind. The world simulation is now internal to the core of the executing processes, providing the "happy path" for activity. Such process models allow for the tracking of every instance of a state machine for the full lifecycle and decision points are explicit. This is not an emergent property but designed into the agent itself, effectively collapsing the functional distinction between the the world model and agent model. Moreover, it allows for learning and process improvement because there is visibility into the dynamic relationship of the interaction between components.

2.4 Perception/understanding in the stream of events

Perception and language capabilities are not seen fundamentally as huge data processing problems, but a matter of discovering, understanding and controlling patterns in event streams. From this, we suggest a wholesale reconceptualization of the 'grounding problem.' To speak a bit metaphorically, the agent/world model is not internally divided such that the active agent requires fixed and stable ground. The etymology of the term "cybernetics" traces back to the Greek word "κυβερνήτης" (kybernētēs), which means "steersman," "governor," or "pilot." The image of the sailor steering a ship, adjusting course from feedback in the environment gives the analogy for controlling systems. But in the thinking now occurring, we have left behind both land and ocean, existing instead the streaming edge. Leahy's imagery is suggestive of the current predicament:

> As never before the edge of the ocean is the edge of an absolute stream: the explosion of the universe itself, the advent of genetic engineering, the beginnings of the practicability of superconductivity, the incipient technology of thinking: there is no land the stream passes through, no same, neither in imitation of the stream to 'pass through', nor to occupy in refusing to imitate the stream, neither to seize upon nor to seize: the banks are identical with the stream for the first time: in the cosmological flood of the stream of consciousness both banks flow as the stream itself flows. [33]

The thinking machine must actualize its very first thought in the stream, not hanging back from the flow in relationship to the flow, which is all current technology can do. Event-driven AGI can be a system that is always beginning, innovating, creating and sharing the absolute edge of consciousness. Philosopher Edward Casey, in an essay on the thinking now occurring, elaborates: “Here, *event* is again a key word—for an absolutely actual *new world* is now eventfulness itself. This actual newness of the world, the likes of which the world has never seen, is the work, the deposition, of an absolute edge unlike any other.” [34] The ever growing activity model integral to the AGI is the creation of this absolute edge, the composite set of these patterns, which are recognized and utilized to achieve specific, often complex, objectives.

2.5 Causality

If the goal is to create AGI, it seems unwise to continue down a path which inherently focuses on capturing the statistical patterns in data rather than understanding the underlying causal relationships. There is no clear path forward for LLMs to reliably solve complex decision-making. The foundational limitation of the data-processing approach for large language models and other AI systems in achieving Artificial General Intelligence is that these models are primarily based on statistical pattern recognition and correlation rather than understanding causality or world modeling. LLMs are trained on vast amounts of data (trillions of tokens) to find patterns and correlations but often fail to capture causal relationships. Understanding causality is critical for achieving AGI, as it enables better reasoning, generalization, and decision making. This capability must be enabled at the level of the programming language, not as an algorithmic and inferential afterthought.

3 The Autonomy Gap

In this section, we reflect on the crucial role of autonomy for artificial general intelligence. As mentioned above, we part ways with Google’s assessment of the place of autonomy in AGI; “The team also points out that AGI does not imply autonomy.” [35] The arc of technological progress is to go beyond mere automation to achieve truly autonomous decision-making systems. In a post-IT world, the so-called “6th wave” which is only now beginning has been deemed “The Era of Autonomy.” [36]

In the past, natural questions to explore were: Which IT tasks can be automated to have the workforce focus on higher-value ones? Or how can IT infrastructure be secured to enable remote working? By integrating the capabilities of DADA X, the conversation evolves. We can now raise the following new question: “Which processes can be made fully autonomous and simultaneously fortified with autonomous security?”

3.1 Defining Autonomy

Autonomy is the capacity to act, make independent decisions, and self-govern without external control. Autonomous systems and applications are the cornerstone of a new technological renaissance and key to AGI: "Autonomous agents have long been recognized as a promising approach to achieving artificial general intelligence (AGI), which is expected to accomplish tasks through self-directed planning and actions." [37]

Distinguished from Automation. Autonomy represents a higher level of technological sophistication than automation, which follows fixed paths. Automated systems execute predefined tasks consistently, without human intervention, but without the capability to learn or adapt to new conditions. Automation is a weak and limited technology with rigid, fixed rules and no understanding of what it executes.

Even in "The Basis of Pragmaticism," written in 1906, C.S. Pierce differentiated the two, but with little optimism about the prospect of progressing beyond automation:

> …it is quite true that we cannot make a machine that will reason as the human mind reasons until we can make a logical machine (logical machines, of course, exist) which shall not only be automatic, which is a comparatively small matter, but which shall be endowed with a genuine power of self-control; and we have as little hopes of doing that as we have of endowing a machine made of inorganic materials with life…if we could endow a system of signs with self-control, there is very strong reason to believe that we should thereby have conferred upon it a consciousness…[38]

Nevertheless, a key point relevant to AI is made, according to Pierre Steiner, that "the difference between human reasoning and machine reasoning is basically related neither to consciousness nor to originality, but to the degrees of control, purpose, and reflexivity human reasoning can exhibit" [39]

The Automation Paradox. Although intended to increase efficiency, automation has instead generated new and unforeseen burdens and amplified risks. This inherent inflexibility is the crux of our escalating cybersecurity and operational vulnerabilities as static systems are prime targets for relentless and evolving digital threats. If one is to speak of an evolution in technology, the next obvious step is to guide that evolution from automation to autonomy. A different approach is needed for this move to a higher level.

Defining Agents. Within the discussions of AGI, autonomy is often a buried concept, either unaddressed or hidden in the idea of agency, or the ability to control future states. To perform human-level cognitive tasks, AGI must not be a mere *tool* like current AI applications, but an *agent*: a program or system that can independently observe its environment and make decisions or take actions to achieve specific goals.

We are familiar with the notion of levels of autonomy from the context of the automobile industry. In 2014, the Society of Automotive Engineers International published the influential J3016 standard which defined six levels of driving automation. The team at Google's DeepMind sees autonomy in terms of similar levels where "higher levels of autonomy are 'unlocked' by AGI capability progression." [40] We argue that autonomy is a fundamental requirement of any system that is purportedly AGI. Our first priority has been to enable the creation of autonomous agents that can operate, correct, and ensure security without humans in the loop, performing tasks across diverse domains just as a human would.

3.2 Current state of Autonomy

There are no "level five" fully autonomous systems in operation globally today, despite many years of pursuit. Ilya Sutskever, chief Scientist of OpenAI, stated in an interview in Nov. 2023: "Right now, those systems are clearly not autonomous; they're inching there, but they're not, and that makes them a lot less useful too…" [41]

Humans still in the loop. Human backup and intervention currently plays an essential role in current vehicle technology that has been marketed to the public as autonomous or on the verge of autonomy. In examining the current state of so-called autonomous systems, it's evident that there are significant challenges and limitations. George Mason University professor Missy Cummings, an expert in autonomous systems, who also served as a safety advisor to the National Highway Traffic Safety Administration (NHTSA), shared insights with CNBC. She noted that despite public perception, it has been an industry standard for humans to be on call, actively monitoring the operations of drones, robotics, and autonomous or semi-autonomous vehicles. [42]

Recently, the GM Cruise Autonomous Taxi program, has had to suspend its driverless operations due to a collision incident. GM disclosed that these autonomous taxis necessitated human intervention roughly every four to five miles. This translates to a significant human presence in the backend operations, with one remote assistant overseeing approximately 15 to 20 vehicles. [43]

Even low level automation is using humans to make up for the problems with current AI systems. The exaggerated marketing of AI agent systems leads tech companies to obscure the true extent of the need for humans in their operation. Another recent example was reported involving order taking automation at some fast food chains. The AI drive thru system relied on outsourced human labor for its effectiveness. Presto Automation, the provider of these drive-thru systems, disclosed in filings with the US Securities and Exchange Commission that it employs off-site agents, particularly in countries like the

Philippines, to assist its "Presto Voice" chatbots. These agents are involved in over 70% of customer interactions, indicating a significant level of human intervention in what is marketed as an automated AI system. [44]

3.3 Challenges for Achieving Autonomy

It should be obvious that Level 5 Autonomy cannot be achieved as long as process operations happen in batch mode. But the problem is not merely due to temporal lag. The barrier is inherent to the methodology being used. Consider: the schema of databases is generated from process maps. But the schema does not include the communication interface information between events, hence crucial causal/contextual relationships are lost. Even if the database were able to preserve the causal relationships (which with our approach is possible), control and correction needs to take place on the fly, before the data is stored. In an autonomous system, events fuel the decision-making engine, and data is reduced to something akin to exhaust.

3.4 Causal tracking vs. causal inference

Our focus extends to leveraging causality in the service of autonomy. Our goal is to develop AGI that can not only understand causal relationships but also use this understanding to make autonomous decisions in real-time. By prioritizing the development of models that understand causality, we can build more robust and trustworthy systems. To perform the needed tracking, the accurate representation of causality, time, and states is crucial.

To anticipate the discussion a bit, and lead into it, we can say that a way of representing causality using events is required. And the most effective way to do this is to have causal operators directly incorporated into the base programming language (in our case, the *Rapide* ADL). This significantly simplifies the process of causal tracking, as the language itself is designed to capture and model cause-and-effect relationships without the need for correlation rules. This is a departure from traditional AI models that require complex algorithms and extensive data in an attempt to infer causal relationships. Filip Piękniewski, a researcher working on computer vision and AI states the problem strongly "This distinction between weak, statistical relationship and a lot stronger, mechanistic, direct, dynamical, causal relationship is really at the core of what in my mind is the fatal weakness in contemporary approach in AI." [45] With AI, you have no choice: you can not track, only infer, because there are no states to be monitored. The challenge with data-centric approaches lies in their focus on mathematics to the exclusion of formal logic and soundness, leading to an incomplete understanding of the interplay between causality, time, and state. These approaches often treat these elements of causality, time and state in isolation, using concepts like Directed Acyclic Graphs (DAGs).

3.5 Decision-making

The relationship between autonomy and decision-making is fundamental, whatever the context. Unless an AGI system can perform live audit on its processes, detecting and making the decisions required to remediate anomalies before they are executed by the system, that system lacks a crucial capability for general intelligence. In approaches where causality is only inferred, decisions require extensive manual auditing.

4 The technology of DADA X

4.1 The *Rapide* Language

The question of which programming language is most appropriate for building AGI is sometimes raised. Python is currently the most popular language used by those building AI applications, although others have been suggested, such as Rust. But what properties are essential for a programming language to enable the cognitive tasks required for an autonomous AGI system? Because causality is the key to sustaining a coherent world model and enabling autonomous capability, the *Rapide* ADL was chosen for building DADA X. It forms the basis for solving a wide variety of problems because it can 'speak the language of causality.' We think this is a missing ingredient in the recipe for AGI.

Rapide is a language specified at Stanford University (Dr. David Luckham's Program analysis and Verification Group) when DARPA requested a language capable of handling complex event processing for distributed systems. It is unique in using causal operators, explicitly representing various kinds of dependencies between events. It offers a unique model of computation based on partially ordered sets of events (event patterns), enabling complex transactions and truly concurrent activities to be managed with accuracy and precision. [46] A pattern serves as a template which can correspond to a partial order of events, provided that the events and their orderings adhere to the pattern's semantic rules. DADA X is the only commercial software platform that uses the *Rapide* language. Its unique ability to track causality and perform backward-checking sets it apart from traditional data-driven approaches that only attempt to infer causal relationships using complex algorithms and extensive data.

4.2 Event-driven Architecture

To speak metaphorically, the fuel for AI is data, but the fuel for AGI will be events. The cornerstone of Decision-Zone's approach is a departure from traditional data-driven processing towards an event-driven paradigm. Because DADA X processes event patterns—complex live data— it must continuously monitor numerous event channels on the message bus. Think of patterns as micro-rule engines at the input side of a system.

The view of Professor Zhu Song-Chun, a leading computer scientist and the director of the Beijing Institute for General Artificial Intelligence, is that our knowledge of objects is task-driven. [47] China's initiative with the Beijing Institute for General Artificial Intelligence is quite telling. They are specifically focusing on causality in their system, and on the task-driven (or event-driven) approach, moving past traditional, data-driven models. They call it the "Small Data, Big Task" paradigm. The recognition of the need for an event-based and causal paradigm is a validation of the approach that Decision-Zone took when conceiving the DADA X platform.

Tasks require states. The concept of "state" in concurrent programs is inherently more complex than in sequential programs. In concurrent programs, multiple processes execute independently, which means a singular, sequential understanding of program state doesn't apply. Divergence from a linear state model requires a more sophisticated approach to state management.

All computable tasks essentially require some form of stateful processing to be done effectively, so employing stateless specifications which do not keep track of states is merely a superficial solution. This leads to suboptimal outcomes, particularly in handling exceptions. This is a fundamental tension in computational problem-solving: the trade-off between the theoretical elegance or simplicity of stateless systems and the practical efficiency of stateful systems.

The value of a platform built using *Rapide* is the ability to work at the state machine level, which brings an incredible amount of operational efficiency. Unlike traditional programming environments that struggle with recursive instances, DADA X efficiently monitors and live audits the entire lifecycle of each state machine which enhances process visibility.

At its core, DADA X operates on an event-driven model, where it listens to events on the message bus and audits for valid patterns. This means that instead of processing raw data, DADA X monitors the sequence and pattern of events happening in real-time. Its event-driven architecture is capable of maintaining and interpreting the state of concurrent processes, crucial for the development and operation of autonomous applications.

Generating event patterns. Recall Goertzel's approach, rooted in a philosophical perspective that emphasizes the importance of patterns in understanding intelligence. This is a key point with which we are in full agreement. DADA X pattern language generation is powerful and efficient, streamlining the process of capturing a wide array of events communicated across the bus, effectively creating a detailed map of all the IDs within the operational environment. The process begins with pattern selection, where patterns can be picked out to be automatically generated. This dramatically simplifies the development process, eliminating the need for extensive manual coding. A task that would typically require hundreds of thousands of lines of manual code is now executed in just a few lines. According to Goertzel: "Efficient pragmatic general intelligence is a measure that rates an agent's intelligence higher if it uses fewer computational resources to do its business." [48] On this measure, DADA X excels beyond all others.

The pattern can also be tailored as needed to suit specific requirements. Our generator offers a builder to ease this process. It generates an expression featuring branches, custom nodes, and exit expressions.

Event Sources. There is no limit on how many event sources can be used. The DADA X platform can handle 100 million binary decisions per second and with capability DADA X could fine tune a system to execute with maximum performance. In addition, 100 million live anomaly decisions per second dramatically reduces cost of ownership.

4.2 The crucial role of Activity Models

> "The brain of a robot and the human brain can both be seen as a model."
> -Prof. Zhu Song-Chun, AI: The Era of Big Integration [49]

An algorithmic data processing logic and a UML (Unified Modeling Language) model logic for business processes serve different purposes and represent different types of information. Unlike AI models, which are data models, DADA X utilizes process models. UML models are used to visualize and describe the flow of activities, interactions, and data within a process. We have abstracted the event layer into a higher layer, the model. The ability to work at the state machine level brings an incredible amount of operational efficiency.

Consider the data layer as akin to a simple, flat structure, much like an Excel spreadsheet with columns. Each column represents different event sources, yet they exist without any inherent connection. The only method to discern the connections between these columns is by examining the rows. In an AI approach, column 1 is interlinked with all other columns through a specific algorithm. This algorithm is derived from analyzing the data spanning from row 1 to row n, resulting in a linear regression formula. When a new row of data is introduced, the model predicts probabilities, but does not inherently reveal the underlying relationships and lacks any mechanism to discern causal relationships. In such systems, columns are merely correlated with time stamps, devoid of any deeper understanding of causality's hierarchy or its underlying logic. The fundamental limits of deep learning and statistical modeling have been noted by many researchers, including Zhu Song-Chun, leading him to argue that we need “fundamental shift in how we think about AI.” [50]

Our models provide a conceptual representation of how all elements are interconnected. This is a top-down approach which makes it easier to visualize and understand the interconnections and dependencies between various steps in an activity, allowing for more efficient development and maintenance of highly scalable event-driven applications.

Such modeling provides a high level of abstraction for the process design, and activities can be defined and modeled to the highest degree of granularity. In a process-centric approach, the system focuses exclusively on monitoring authorized patterns of process events. Consequently, any events that deviate from these authorized patterns are identified as violations in the operational process. One advantage of a process-centric approach is that the number of authorized process event patterns to be monitored and mediated is generally a smaller, more manageable number.

Implications for Development. The comprehensive toolkit provided allows for the importation of UML state machines, their conversion into XMI files, and their subsequent direct integration into the system. This system is designed to automatically determine the appropriate whitelist for the state machine, thereby removing the need for manual rule writing. It is capable of effectively segregating the business logic.

The decentralized autonomous decisioning agent fundamentally changes the development landscape by allowing developers to go from models directly to implementation with as-is logic. This approach eliminates the need for manual coding and manual testing, saving time and reducing the risk of human error. Changes to a system can be made at the model level. Model-driven development enables the verification of the model itself, the validation of the system in accordance with that model, and the execution of simulations to confirm the accurate functioning of the model.

It has been suggested by researchers that multiple LLMs can be used as distinct “experts” for specific sub-tasks. [51] DADA X uses discovered or predefined behavior models and LLMs such as Bard can be a source for needed models as they are capable of generating various processes such as an activity diagram for the life cycle of loan financing, or for an automobile braking system. In addition, they can be asked to specify the associated event providers.

DADA X reduces the development lifecycle from years to months. Perhaps it goes without saying that this method of designing the software is substantially more cost-effective. Ultimately, organizations can significantly reduce both development and maintenance costs. Clearly, there is a huge benefit to eliminating patching cycles which often require downtime and require significant resources.

Metacognition and anomalies. Although not specifically addressed earlier in the discussion of the cognitive tasks required for general intelligence, metacognition is often mentioned as well. This is a higher level of understanding that is sometimes described as reflection, or thinking about thinking. On the human level, a comparison can be made to the notion of auditing (in a business or technical context) which is a systematic process of evaluation and analysis. But more generally we can say that this identification of anomaly is the key to initiating the subsequent process of reflection and corrective action that is thereby made possible.

When auditing to ensure the verifiability, integrity and accuracy of a process or information, the ability to detect anomalies is the difference between a reflective system and a purely automatic or unreflective one. With respect to human existence, as the philosopher Martin Heidegger elaborated phenomenologically, reflection is often the result of a breakdown in one sense or another. In work and life, when everything is running smoothly we do not need to stop and reflect. It is when the tool is missing or defective and a problem arises that the flow is broken and the task changes to become one of correction and remediation. [52] We find a similar view in Peirce, for whom we are naturally inclined towards a state of mental repose (or a default state of not thinking) from which we begin to engage in thought as a result of some surprise or unexpected event that challenges or disrupts our existing understanding. Active thinking itself is just the attempt to resolve the inconsistency that has been introduced when we are confronted by the concrete surprise so that we can go back, to not thinking.

Currently no software system in the world, apart from DADA X, can make a decision on a single event as to its correctness. We consider this ability to be a sort of "Holy Grail" of information technology. Operating on the front end means near zero latency for live auditing, control, and security at the input level, preventing the execution of anomalies by the system. All other software systems require exception management which represents more than 80 percent of the code. There are billions of exceptions in any given software and only a few thousand exceptions are actually coded. A system that is restricted to determining after the fact whether an anomaly has occurred, even if it could self-code exception rules to handle future similar exceptions, is not yet at human-level intelligence.

What this illustrates, as we find in the philosophy of world pragmatism, is the important observation that now, thought itself surmounts the barriers and rises "above the rule." Because only the correct process is monitored, exception rules are no longer required. In Leahy's terms, the fact itself is synthesized. For DADA X, the event, at its most granular level, is completely intelligible. According to Leahy, "The matter of the thing is the absolutely formal being of thought, that the thing in essence is nothing but the appearance itself." [53] In the context of the thinking machine, this translates to the agent's ability to directly grasp the essence of events by analyzing their patterns and relationships as they unfold. The models represent the agent's "formal being of thought," allowing it to identify anomalies and act to remediate or to control the system. DADA X transcends the role of a mere tool and becomes an intelligent operator where the intelligent behavior is realized in directly managing the flow of events dynamically in real-time, scaling as the complexity increases and managing all instances without the problem of recursivity.

In essence, DADA X acts as a real-time, autonomous decisioning agent capable of interpreting context, understanding causal relationships, and making instantaneous decisions about work activity instances.

6 Learning

We take a process-centric and "acquisitive" approach to learning, going against the grain of current trends in big data driven deep learning. The key cognitive algorithm we use is our proprietary discovery algorithm. A key difference with our AI is that 'training' occurs on real-time events, not log data. The events are modeled by this algorithm and UML models are generated. This proprietary causality and standard AI algorithm dynamically extracts baseline behavior process models that capture the logic of a system and build state machines. This is essentially reverse engineering, using events with AI.

7 Security

The security of an AGI system will be as important as that of any other critical infrastructure, if not more so. Current software technology cannot protect systems autonomously. Traditional approaches operate at audit-time, deciding after the fact if an anomaly has occurred. AGI must autonomously protect itself, replacing human intervention in the dynamic identification and neutralization of incidents and intrusion attempts.

There are particular and significant risks facing custom GPTs due to their weak security: "Through prompt injection, an adversary can not only extract the customized system prompts but also access the uploaded files." Researchers systematically demonstrated the vulnerability, finding that their prompt injection attacks on custom GPTs were suprisingly successful (a 97.2% success rate for system prompt extraction and a 100% success rate for file leakage). [54]

Our architecture protects the input layer. Decentralized autonomous security monitors and mediates authorized process events, providing cybersecurity at machine speed by preventing process violations from occurring, rather than responding after the fact. A front-end approach monitors the message bus for the correct order of operations, detecting and remediating potential security breaches in real time, which is a significant improvement over traditional historical, database-driven approaches to security management. Cybersecurity is not an add-on to the platform, but is at the core of the model-driven approach.

The transformation from audit-time decision making to real-time monitoring and autonomous remediation is critical for managing increasing systemic management and security issues. Anomalous messages can instantly trigger remediation applications to prevent catastrophic breaches. This autonomous sense and response capability is critical for handling threats within milliseconds, providing a high accuracy of threat detection, and pinpointing the source of threats, thereby addressing the attribution problem in cybercrime.

8 Ethical Considerations

"Nobody knows what's going on in there."
-Eliezer Yudkowsky [55]

Black Box Risk. Those who are most vocal and vehement about the possible risks associated with AGI tend to focus on concerns surrounding agency. But these concerns are deepened by the fact that the current discourse is focused on a vision of AGI which is an extension of current "black box" approaches in AI which are characterized by opaqueness and lack of interpretability. Deep learning models can be challenging to interpret, making it difficult to understand how a model is making its predictions. This can be a significant challenge in applications where transparency and accountability are crucial, and certainly the case with AGI systems as well. If users, developers, or regulators cannot understand how an intelligent system arrives at its conclusions or predictions, this lack of explainability leads to issues with trust when such systems are deployed.

Using our platform, applications can be designed to reliably pursue human-aligned goals. This is especially crucial for AGI systems that attain higher levels of autonomy, as it's imperative to ensure that these agents act beneficially towards humanity and do not inadvertently cause harm. Moreover, the platform uniquely offers transparent understanding and justification for each decision, bridging the often daunting interpretability gap associated with AGI. Since DADA X live audits every event and produces the audit trails, "interpretation" is no longer an issue. We know what's going on in there, at every level of abstraction in the model hierarchy. This type of model-based platform offers a clearer path towards AGI alignment.

Alignment. Open AI leaders have proposed steps for navigating the development of superintelligent AI, including the idea of using human level AGI to align an ASI. As many have pointed out, individual humans and societies not only have conflicting and competing values, but are often internally divided such that they fail to achieve alignment with their own ethical or moral ideals. One worry recently expressed by Elon Musk is that a future AGI can be polluted if trained on data scraped from the internet which puts forth evil values. What those are, specifically, can be disputed, but in any case, garbage in garbage out is an adage that holds true. The remarkable technology of the transformer architecture which enabled the current explosion of generative AI began with the paper: "Attention is all you need." But the data sets for training the neural nets are carefully curated by an army of trainers, which is also the way biases can enter the model. Reinforcement with human feedback is used and so-called guardrails are then applied to try to keep the outputs of these models in check. Leahy, speaking in an ethical vein, states: "The complex simplicity of beginning requires attention. What happens does so neither automatically nor according to plan." [56]

For actual good— beneficence as opposed to mere benevolence— readiness and discretion must be actual as well. Readiness is defined as 'being at the disposal of another.' Discretion is defined as 'having the patience to see what's different.' "The required attention to the complex simplicity of beginning precludes forcing." [55] The system not be forced or coerced but rather prepared to attend to the correct process and discriminate/discern. If action is taken without readiness and discretion, capabilities are misused. Proper attention to modeling enables the readiness and the discretion of an ethical AGI. A fuller treatment of the ethical and alignment issues related to DADA X platform for AGI will be the concern of a future paper.

9 Future Directions

While our core technology is robust, we recognize that research and development are vital for building applications on our platform and unlocking its potential.

Although our primary focus here has been the DADA X platform with its causal tracking and autonomous capabilities, this "engine" for AGI will require increasingly more and complex models. At first, subject matter experts may provide such models. But for the system to learn the discovery and generation of these models will eventually become autonomous as well. It is here that learning takes place, at the edge.

The recipe for AGI using the technology described above is modular and event-driven. Events are streamed to the learning/discovery algorithm, which then generates models and patterns in order to reverse engineer the underlying world model. The result is the creation of a massive pattern base, or Large Process Model (LPM), which is at a higher contextual level than a language token database. This effectively replaces the neural net and transformer architecture. This is linked directly to the DADA X. Natural language inputs to the system can then be used to generate UML and patterns for execution, testing, and validation. As more models and patterns are generated the Large Process Model, grows.

True Decentralization. We can envision future AGI as a distributed synthetic intelligence where the Discovery algorithm generates state machines which are then deployed as models into DADA X agents to control various applications and aspect of integrated systems. Autonomy moves to new ground, what Leahy has referred to as "metonomy," with the synthesis of model generation and model control.

AGI On device. AGI systems will require the ability to make decisions, learn, see and understand speech. Our long term goal is to create a four chip set for Artificial General Intelligence: (1) the DADA X Agency/Autonomy Chip (2) Discovery AI Algorithm for reverse engineering of process models (3) Natural language acquisition and (4) event-driven Vision unit. DADA X intelligence local to devices will reduce latency and provide autonomous capabilities even in demanding environments. This is the genesis of a new era in AI, where machines adapt dynamically, and navigate our world with a level of autonomy and intelligence previously unimagined. Our chips for on-device AGI will make this possible. We envision a world where autonomous, event-driven applications dominate, fundamentally altering the landscape of computing, business operations, and cybersecurity.

10 Summary

The growing complexity of data generation and systems management is stretching human capabilities and resources to their limit. The world is eager for mature, transparent, and dependable AI technology, but current data-driven methods are struggling with inherent limitations.

Key insights are that AGI and related future applications should work with *events,* not data, and must *decide first and then act, in real-time*, not audit time. The world as it is actually experienced by human beings is one of a continual stream of events in consciousness. Event-based AGI can work in the present moment, observing, understanding, responding and controlling the live stream of complex events passing into the system. These factors have greater direct potential to effect 'machine thinking' than next-word token prediction.

In summary, the unique capabilities of the DADA X platform are based in its more expressive programming language which uses causal operators, and a sophisticated event-driven architecture to monitor, analyze, and react in real time, based on a consistent understanding of events and their causal relationships. This enables the creation of more robust, secure, and efficient autonomous systems.

References

[1] Luciano Floridi "What is the Philosophy of Information?" *Metaphilosophy*, 2002, (33), 1/2.

[2] Meredith Ringel Morris, Jascha Sohl-dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet and Shane Legg "Levels of AGI: Operationalizing Progress on the Path to AGI" November 4, 2023 p.5 URL https://arxiv.org/pdf/2311.02462.pdf

[3] Irene Zhang "AI Proposals at 'Two Sessions': AGI as 'Two Bombs, One Satellite'?" March 8, 2023 URL https://www.chinatalk.media/p/ai-proposals-at-two-sessions-agi

[4] Meredith Ringel Morris, Jascha Sohl-dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet and Shane Legg "Levels of AGI: Operationalizing Progress on the Path to AGI" November 4, 2023 p.4 URL https://arxiv.org/pdf/2311.02462.pdf

[5] Dwarkesh Patel "Shane Legg (DeepMind Founder) - 2028 AGI, Superhuman Alignment, New Architectures", October 26, 2023 URL https://www.dwarkeshpatel.com/p/shane-legg

[6] Will Douglas Heaven "Google DeepMind wants to define what counts as artificial general intelligence" MIT Technology Review Nov. 16, 2023 URL https://www.technologyreview.com/2023/11/16/1083498/google-deepmind-what-is-artificial-general-intelligence-agi/

[7] https://openai.com/charter Published April 9, 2018

[8] Ben Goetzel "The General Theory of General Intelligence: A Pragmatic Patternist Perspective" April 6, 2021, p.9.

[9] Ben Goetzel "The General Theory of General Intelligence: A Pragmatic Patternist Perspective" April 6, 2021, p.39.

[10] Joscha Bach "From Large Language Models to General Artificial Intelligence" URL https://www.youtube.com/watch?v=9IBHkY28D5g

[11] Gary Marcus URL https://twitter.com/GaryMarcus/status/1727887582871306360

[12] Peter Voss "What is AGI? Feb 21, 2017 URL https://medium.com/intuitionmachine/what-is-agi-99cdb671c88e

[13] Meredith Ringel Morris, Jascha Sohl-dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet and Shane Legg "Levels of AGI: Operationalizing Progress on the Path to AGI" November 4, 2023 p.9 URL https://arxiv.org/pdf/2311.02462.pdf

[14] Sébastien Bubeck et al, "Sparks of Artificial General Intelligence: Early Experiments with GPT4" April 13, 2023, arXiv:2303.12712v5

[15] Video interview with Dwarkesh Patel "Ilya Sutskever (OpenAI Chief Scientist) - Building AGI, Alignment, Spies, Microsoft, & Enlightenment" URL https://youtu.be/Yf1o0TQzry8?si=Gf-X_R7Sma9X-UuQ

[16] Mark van Rijmenam "How To Trust Each Other in A Post-Truth World" Jan. 2, 2024 URL https://www.thedigitalspeaker.com/how-to-trust-each-other-in-a-post-truth-world/

[17] Video interview "Joscha Bach and Connor Leahy" June 19, 2023URL https://youtube.com/watch?v=Z02Obj8j6FQ&t=801s

[18] Herbert L. Roitblat "The science of (artificial) intelligence" November 15, 2023 URL https://bdtechtalks.com/2023/11/15/the-science-of-artificial-intelligence/

[19] Rylan Schaeffer, Brando Miranda, and Sanmi Koyejo. "Are Emergent Abilities of Large Language Models a Mirage?" arXiv preprint arXiv:2304.15004, 2023.

[20] Pierre Steiner "C.S. Peirce and Artificial Intelligence: Historical Heritage and (New) Theoretical Stakes" V.C. Müller (Ed.): Philosophy and Theory of Artificial Intelligence, SAPERE 5, p. 266.

[21] Stephen Hoath in Lissa McCullough and Elliot R. Wolfson, eds. D. G. Leahy and the Thinking Now Occurring State University of New York Press, 2021. p.271-272.

[22] David G. Leahy *Foundation: Matter the Body Itself*. Albany: State University of New York Press, 1996. p.ix.

[23] Pierre Steiner "C.S. Peirce and Artificial Intelligence: Historical Heritage and (New) Theoretical Stakes" V.C. Müller (Ed.): Philosophy and Theory of Artificial Intelligence, SAPERE 5, p. 266.

[24] Lissa McCullough and Elliot R. Wolfson, eds. *D. G. Leahy and the Thinking Now Occurring* State University of New York Press, 2021. p.5.

[25] David G. Leahy *Foundation: Matter the Body Itself*. Albany: State University of New York Press, 1996. p. 116.

[26] Charles S. Peirce, *The Essential Peirce: Selected Philosophical Writings (1893–1913), Vol. 2.* Edited by Nathan Houser and Christian J. W. Kloesel. Bloomington, IN: Indiana University Press, 1998. 2.356

[27] *The Collected Papers of Charles Sanders Peirce*, reproducing *Vols. I-VI*, Charles Hartshorne and Paul Weiss, eds., 1931-1935, Harvard University Press, Cambridge, Mass., and Arthur W. Burks, ed., 1958, *Vols. VII-VIII*, Harvard University Press, Cambridge, Mass. CP 8.332 URL https://colorysemiotica.files.wordpress.com/2014/08/peirce-collectedpapers.pdf

[28] David G. Leahy *Foundation: Matter the Body Itself*. Albany: State University of New York Press, 1996. p.460.

[29] David G. Leahy *Foundation: Matter the Body Itself*. Albany: State University of New York Press, 1996. p.459.

[30] David G. Leahy *Foundation: Matter the Body Itself*. Albany: State University of New York Press, 1996. p.461.

[31] David G. Leahy *Foundation: Matter the Body Itself*. Albany: State University of New York Press, 1996. p.462.

[32] David G. Leahy *Beyond Sovereignty: A New Global Ethics and Morality*. Aurora, CO: The Davies Group, 2010. p.241.

[33] David G. Leahy *Foundation: Matter the Body Itself*. Albany: State University of New York Press, 1996. p.591.

[34] Edward Casey in Lissa McCullough and Elliot R. Wolfson, eds. *D. G. Leahy and the Thinking Now Occurring* State University of New York Press, 2021. p.247.

[35] Meredith Ringel Morris, Jascha Sohl-dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet and Shane Legg "Levels of AGI: Operationalizing Progress on the Path to AGI" November 4, 2023 p.13 URL https://arxiv.org/pdf/2311.02462.pdf

[36] John Cowan "Public-Private Partnerships and the Era of Autonomy" June 7, 2023 URL https://nextwave.partners/insights/public-private-partnerships-and-the-era-of-autonomy

[37] Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, et al. A survey on large language model based autonomous agents. *arXiv preprint arXiv:2308.11432*, 2023.

[38] Charles S. Peirce, *The Essential Peirce: Selected Philosophical Writings (1893–1913), Vol. 2*. Edited by Nathan Houser and Christian J. W. Kloesel. Bloomington, IN: Indiana University Press, 1998. 2.387

[39] Pierre Steiner "C.S. Peirce and Artificial Intelligence: Historical Heritage and (New) Theoretical Stakes" *V.C. Müller (Ed.): Philosophy and Theory of Artificial Intelligence*, SAPERE 5, p. 274.

[40] Meredith Ringel Morris, Jascha Sohl-dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet and Shane Legg "Levels of AGI: Operationalizing Progress on the Path to AGI" November 4, 2023, p.10 URL https://arxiv.org/pdf/2311.02462.pdf

[41] No Priors Ep. 39 | With OpenAI Co-Founder & Chief Scientist Ilya Sutskever URL https://www.youtube.com/watch?v=Ft0gTO2K85A

[42] URL https://www.cnbc.com/2023/11/06/cruise-confirms-robotaxis-rely-on-human-assistance-every-4-to-5-miles.html

[43] URL https://www.cnbc.com/2023/11/06/cruise-confirms-robotaxis-rely-on-human-assistance-every-4-to-5-miles.html

[44] Frank Landymore "Busted! Drive-Thru Run by 'AI' Actually Operated by Humans in the Philippines" December 8, 2023 URL https://futurism.com/the-byte/drive-thru-ai-humans

[45] Filip Pieknewski "Ai Reflections" April 9, 2023 URL https://blog.piekniewski.info/2023/04/09/ai-reflections/

[46] “Rapide Language Tutorial” URL https://complexevents.com/stanford/rapide/examples/teaching/dtp/language.html

[47] Zhu Song-Chu “AI: The Era of Big Integration Unifying Disciplines within Artificial Intelligence” p.18 URL https://dm.ai/wp-content/uploads/ebook-ai-the-era-of-big-integration.pdf

[48] Ben Goetzel “The General Theory of General Intelligence: A Pragmatic Patternist Perspective” April 6, 2021, p.26.

[49] Zhu Song-Chu “AI: The Era of Big Integration Unifying Disciplines within Artificial Intelligence” p.17 URL https://dm.ai/wp-content/uploads/ebook-ai-the-era-of-big-integration.pdf

[50] Zhu Song-Chu “AI: The Era of Big Integration Unifying Disciplines within Artificial Intelligence” p.18 URL https://dm.ai/wp-content/uploads/ebook-ai-the-era-of-big-integration.pdf

[51] Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, et al. A survey on large language model based autonomous agents. *arXiv preprint arXiv:2308.11432*, 2023.

[52] Heidegger, M., *Being and Time*, (trans., Macquarrie and Robinson) New York: Harper and Row, 1962.

[53] David G. Leahy *Foundation: Matter the Body Itself*. Albany: State University of New York Press, 1996. pp.5-6.

[54] Jiahao Yu, Yuhang Wu, Dong Shu, Mingyu Jin, and Xingyu Xing. Assessing Prompt Injection Risks in 200+ Custom GPTs URL https://ar5iv.labs.arxiv.org/html/2311.11538

[55] Eliezer Yudkowsky: Is consciousness trapped inside GPT-4? | Lex Fridman Podcast Clips URL https://www.youtube.com/watch?v=qplLNwboQM0

[56] David G. Leahy *Beyond Sovereignty: A New Global Ethics and Morality*. Aurora, CO: The Davies Group, 2010. p.88.