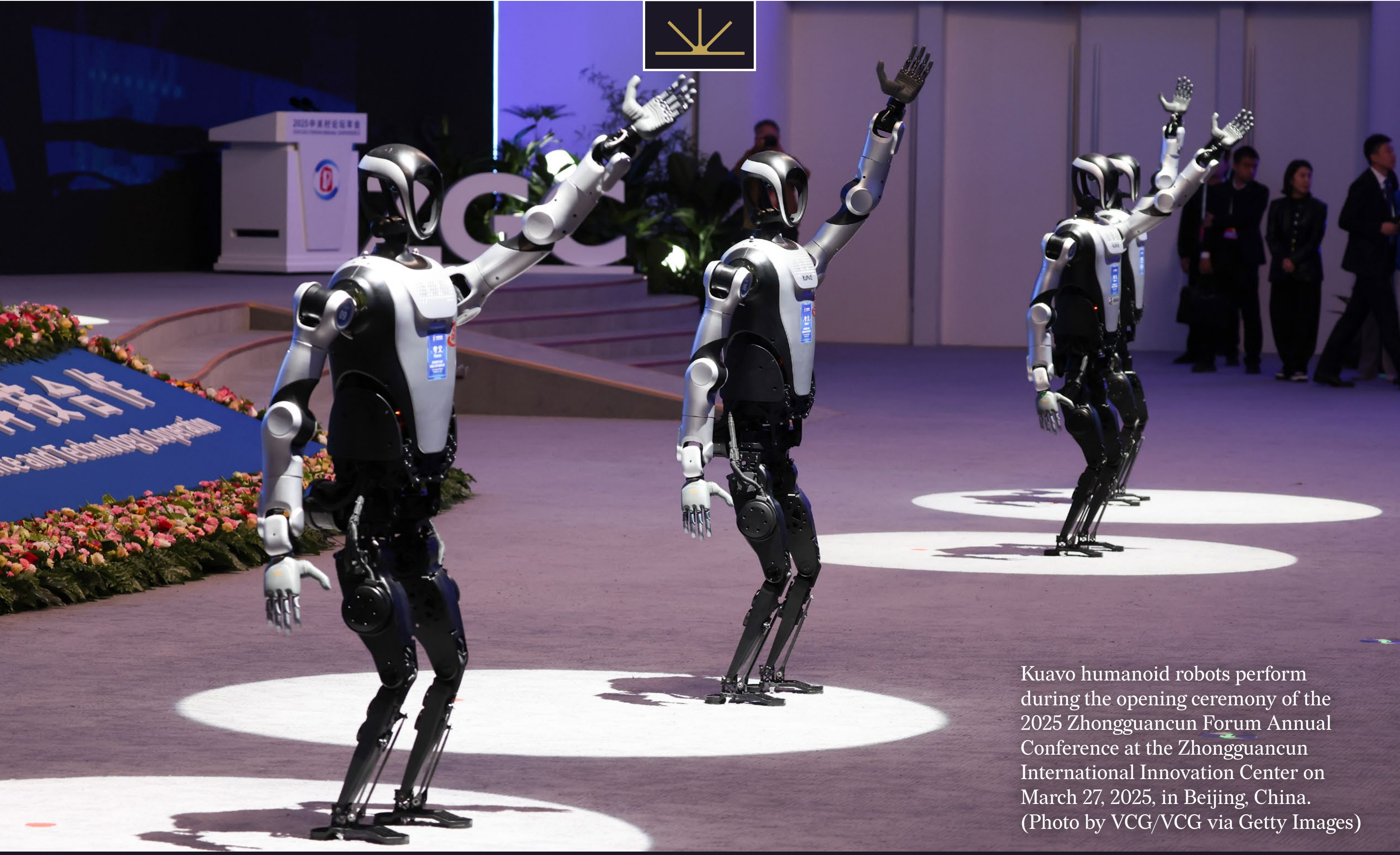


Kuavo humanoid robots perform during the opening ceremony of the 2025 Zhongguancun Forum Annual Conference at the Zhongguancun International Innovation Center on March 27, 2025, in Beijing, China. (Photo by VCG/VCG via Getty Images)

American AI Leadership Should Not Be Defined By Machine Learning

Vincent J. Carchidi

Introduction

U.S. leadership in the development of artificial intelligence should not be defined just by machine learning. This paradigm, in which artificial neural networks learn via data, is a critical step in the progression of this technology. Yet, machine learning is one fundamentally limited paradigm whose shortcomings cannot be overcome by doubling down on its incumbent techniques. U.S. policymakers should instead reconceive of American AI leadership as investing in and pushing the boundaries of the next dominant paradigm in AI. Neuro-symbolic AI,[1] an emerging paradigm that synthesizes techniques from traditional and contemporary approaches to AI research, is the ideal candidate in this respect. It demonstrates the most promising path to ameliorating shortcomings in state-of-the-art models without sacrificing what came before.

However, U.S. officials increasingly define AI leadership in reference to the material needs of machine learning, namely data, computing power, and energy. In July 2024, then-U.S. Secretary of Commerce Gina Raimondo claimed the superiority of American firms' AI models "wouldn't be the case" were U.S. export controls limiting shipments of advanced semiconductors to China not in place.[2] Her claim implies that the linchpin of American firms' AI leadership is the relative accessibility of computing power.[3]

Raimondo's remark reflects a now-common understanding among policymakers: The path to advanced AI systems is through scale. Scaling up the sizes of models and their training datasets – and then shifting the "scaling up" burden to the time during which models generate outputs – achieves capabilities once the exclusive preserve of human beings. Securing the necessary computing power[4] and energy[5] to train these models is merely the cost of entry, itself the gateway to the "Deep Learning Revolution,"[6] culminating in generative AI.[7] So ingrained is this understanding that the U.S.-China Economic and Security Review Commission's 2024 report recommends establishing a Manhattan Project-like program for "artificial general intelligence" (AGI), complete with executive branch authority to fund multiyear contracts with AI, cloud, and data center firms.[8] This is echoed in analyst recommendations to establish a national computational reserve[9] and to create an AGI commission that helps businesses access data, energy, and computing resources.[10]

This is not the first time that the U.S. government and the AI industry have faced what appears to be the threshold of intelligent machinery. In the early 1980s, the AI systems that captured imaginations were not neural networks but expert systems underpinned by symbolic AI. Their seemingly inexorable rise propelled the Defense Advanced Research Projects Agency (DARPA) to establish the Strategic Computing Initiative in 1983,[11] backed by the U.S. Congress, with the goal of creating a generic expert system capable of underpinning multiple defense applications. While progress in narrow applications was made, the dream of a general system was never realized.[12]

A perceived inevitability accompanies machine learning today. Yet, policymakers have the advantage of hindsight, and with it, a picture of AI's history crystallizes: Advancements made with a new or newly accessible technique in certain areas leads decision-makers to quickly perceive that the achievement of human-level intelligence through this technique is merely a matter of time, only to later realize that the reality of progress in intelligent machinery is never quite as good as it seems.[13]

Policymaking efforts to retain and expand American AI leadership should not concede the future of this technology merely to control its present because its present is fundamentally limited. Machine learning is not the paradigm that will, once fully realized, secure for the U.S. an enduring leadership position in AI. A new paradigm is needed: neuro-symbolic AI. Rather than repeat the mistakes of the past, the U.S. government's role should be relatively targeted and complementary, prioritizing shortcomings in state-of-the-art machine learning systems ripe for improvement in the next paradigm. Rather than pursue AGI, the federal government should invest in frontier neuro-symbolic AI research by laying its foundations through existing offices and programs like the National Artificial Intelligence Initiative Office (NAIIO) and the National Science Foundation's (NSF) National AI Research Institutes.

To make this case, policymakers must understand not only what is happening within the AI industry but what has happened. To that end, a mixed historical, technical, and geopolitical – but accessible – analysis of AI's evolution is provided.

Now is the time to make this argument, as the generative AI boom accounted for over one-quarter of global AI-related private investments in 2023,[14] with U.S. private AI-related investment coming in at a world-leading $67.2 billion in 2023 (compared to China at $7.8 billion).[15] It also comes as U.S. federal funding for AI research and development has more than tripled since fiscal year 2018, with government agencies allocating a total of $1.8 billion in 2023.[16] The highest AI R&D agency requests for FY 2024 came from the NSF ($531 million), DARPA ($322.1 million), and the National Institutes of Health (NIH) ($284.5 million).[17]

Four recommendations are provided, implicating both the U.S. Congress and the Executive Branch:

1. The NAIIO should direct the federal AI Research & Development Interagency Working Group to prioritize long-term investments in neuro-symbolic AI as part of its mandate to promote U.S. AI leadership.

The Godfather of AI' Geoffrey Hinton, speaks at 'Can we control AI?' panel during day two of Collision 2024 in Toronto, Ontario, on June 19, 2024. (Mert Alper Dervis / Anadolu via Getty Images)

2. The NSF should expand its network of National AI Research Institutes by establishing an Institute dedicated to foundational and use-inspired neuro-symbolic AI research in a critical sector, complete with corporate and academic public-private partnerships.
3. Congress should fulfill the promise of the CHIPS and Science Act by increasing federal agencies' basic research budgets.
4. Congress and the Commerce Department should adopt proactive yet targeted export controls on hardware and models in coordination with partners and allies that are proportional to the capabilities of AI models.

The First Two Waves of AI

Fortunately, DARPA adopted a useful conceptualization [18] of AI's technical trajectory to guide understanding of this technology's development, dividing it into three stages: First Wave; Second Wave; and an anticipated Third Wave that ameliorates the shortcomings of the first two.[19] The First Wave was dominated by symbolic AI, characterized by systems built with human knowledge encoded directly into the systems. The current Second Wave is dominated by machine learning, in which neural networks learn via data.

A historical analysis reveals not only what went wrong in the First Wave but why it went wrong. Concomitantly, it shows us how critical the U.S. government's support for basic AI research was before the ascendence of symbolic-based expert systems led to over-promise and under-delivery. The lessons of this history bear directly on the U.S. government's role in the Third Wave.

Foundations of the First Wave

The U.S. government's role in supporting AI as it grew from a collection of scattered research efforts to a recognizable discipline is critical and often overlooked. Although AI originated in the private sector, its early growth was principally dependent on public investments in fundamental research programs. ARPA (later rebranded DARPA) was disproportionately responsible for this transformation through the 1960s to the 1990s, with the initial 10-15 years of AI funding enabling basic and interdisciplinary research without concern for immediate applications. Over time, additional major sources of federal support included other Department of Defense agencies, NIH, NSF, and NASA.[20]

Early pioneers in AI included mathematician Claude Shannon, computer scientist John McCarthy, and then-graduate student Marvin Minsky, who was recruited to work with Shannon and McCarthy at Bell Laboratories. IBM's Nathaniel Rochester shared their belief that AI showed significant promise, with Rochester joining the 1956 Dartmouth workshop on AI.[21] The workshop's associated research proposal[22] is considered a founding document in AI, with all four individuals as coauthors.

That same year, the U.S. Air Force (through Project RAND) funded nearly the entirety of Herbert Simon and Allen Newell's work on Logic Theorist, a computer program that could prove select mathematical theorems. Newell went to work at Carnegie Tech (now Carnegie Mellon University), where the Air Force and Office of Naval Research largely funded the projects on decision-making and problem-solving until the early 1960s. At the Massachusetts Institute of Technology, Minsky and McCarthy established the Artificial Intelligence Project in 1957. Here too, military funding was critical, though informally leveraged through an arrangement with the Research Laboratory of Electronics.[23]

Moreover, ARPA's Information Processing Techniques Office (IPTO) increased funding for Stanford University in 1965 to upgrade computing capabilities, following McCarthy's establishment of the Stanford Artificial Intelligence Laboratory in 1963.[24] Stanford Research Institute's Artificial Intelligence Center, founded in 1966, worked on automatons that could gather, process, and transmit data in a hostile environment, leading to the AI-enabled robot "Shakey," whose construction required basic research in planning, natural language processing, and computer vision. Funders, however, were not satisfied despite progress,[25] foreshadowing the field's perennial discontents.

Such discontents were magnified by external scrutiny in the mid-1970s, leading DARPA Director George Heilmeier (taking office in 1975) to cut the agency's speech understanding research and become more insistent that AI research be linked to mission-oriented applications.[26]

The First Wave's Zenith: Symbolic AI

Throughout this period, artificial neural networks (ANNs) existed, but they were overshadowed by the approach that dominated the First Wave: symbolic AI.[27] Symbolic systems are hand-coded with human knowledge. Think of this feature of symbolic systems as their defining characteristic, separating them from other types of AI. These systems represent human-defined knowledge using symbols, such as words, rules, or formal logic. They do not learn this knowledge from data. The system manipulates these symbols to ascertain relationships between them via the rules or logical statements with which it is innately endowed. Symbolic AI systems thus produce human-interpretable results.[28]

The premise of symbolic AI was simple: once a machine is given sufficient structured facts about the world (handcrafted knowledge), dynamic intelligence will eventually result. Too simple, in fact, as this approach – while finding some successes in expert systems – crashed in the late-1980s as cheap, accessible computers supplanted symbolic-based expert systems that required specialized hardware.[29]

In the late-1970s, however, momentum was shifting towards expert systems. IPTO Director Robert Kahn, who took office in 1979, broke with Heilmeier in seeing real-world promise for them. Simultaneously, there was increasing congressional concern about the threat posed by the Japanese Fifth Generation Computer Systems program to U.S. technological leadership.[30] The result was the establishment of DARPA's Strategic Computing program.

Kahn sold his vision to Congress through the promise of specific AI-enabled applications. Interestingly, the Japanese program powerfully influenced congressional deliberation over Kahn's requests. This partly owed to the publication of computer scientists Edward Feigenbaum and Pamela McCorduck's book "The Fifth Generation,"[31] in which they call on the U.S. government to meet the Japanese challenge. They invoked the idea, as Colin Garvey summarizes, that "expert systems were transforming computers from "calculating machines" that relied on data, to "reasoning machines" that relied on knowledge."[32] Congress approved Kahn's plan by splitting Strategic Computing into two major projects: one for specific applications and another for basic research in support of those applications.[33]

Research in areas like speech understanding that had previously been cut resumed on a large scale in 1984 and persisted into the 1990s with participation from private actors including Carnegie Mellon, MIT, and IBM. DARPA sought to mutually hammer out performance evaluation benchmarks between DARPA managers and funded researchers.[34] DARPA and the National Bureau of Standards (now the National Institute of Standards and Technology (NIST)) held annual system evaluations with government contractor, industry, and university participation. This enhanced rates of adoption and commercialization, though it may have moved away from basic research, as the increased adoption simultaneously lowered the need for basic research.[35]

Unfortunately, early optimism in the potential of expert systems was unwarranted. The Strategic Computing program's founders disagreed on how to best direct its research projects. As Emma Salisbury details, the division was between Kahn, who believed that applications would flow naturally from a developed technology base, and DARPA Director Robert Cooper, who believed that specific applications would give way to a more developed technology base.[36]

This division foreshadowed the downfall of Strategic Computing. By 1988, the program's ambitions were downgraded. By 1993, it was a memory. As Salisbury observed, it is not surprising that an initiative of such interdependent ambition failed. While it did produce successes in computer vision, natural language understanding, and speech recognition, the program failed, she argues, because of overpromise and underdelivery. Indeed, the unusual structure of the program positioned it for this outcome: It funded not only specific research problems but also a multifront, field-wide agenda in which progress in one area was expected to aid the progress of another. This, in turn, was enabled by the founders' willingness to see general AI as a realistic possibility through advancements in computing power.[37]

The Second Wave (Approx. 2012 – present)

The First Wave did not see enduring success. Handcrafted knowledge was not sufficient for the lofty goals of AI, and symbolic AI fell out of favor. The fall of Strategic Computing and the disappointment of expert systems exemplify a familiar boom-and-bust cycle in AI's history, which reverberates in the Second Wave.

The Second Wave eschews the First Wave's reliance on handcrafted knowledge. Instead, it promotes ANNs that learn via statistical associations of data. The ability to learn via data is what separates ANNs from symbolic systems.

This approach is known as machine learning. Its premise is that the assemblies of neurons in biological brains, with all their marvelous interactivity, can be replicated through these artificial networks. A neural network generates predictions about a given task (e.g., predicting the next word, predicting the type of object in an image, etc.). Since the network is not generating these predictions by manipulating human-defined symbols, their predictions are based on their training data. ANNs learn when a word is used or what an object looks like based on that word's or object's distribution in their training data.

Early ANNs were shallow, consisting of a single layer of neurons between input and output. Deep learning simply refers to later ANNs that contain many layers of neurons. Even here, the U.S. government's footprint is visible. One of the first practical instantiations of ANNs – Frank Rosenblatt's Mark I Perceptron[38] – resulted from image recognition work[39] funded by U.S. defense agencies amid broader efforts to develop automatic target recognition.[40] Deep learning pioneer and Nobel laureate Geoffrey Hinton[41] was also supported by NSF funding in the 1980s.[42]

Nevertheless, the U.S. government's role in the Second Wave is markedly different than its role in the First Wave. It is a cliché that machine learning blossomed in the private sector.[43] Companies like Google, DeepMind, OpenAI, Microsoft, and Meta are responsible for the most recent innovations. Indeed, the U.S. government's distanced role in the Second Wave reflects a broader trend. Zegart details how the peak of federal funding for research generally as a share of GDP came in 1964, when it was at 1.9%. By 2020, it had fallen to 0.7%. Basic research funding through major sponsors like the NIH and NSF has struggled, particularly as the latter's budget was cut by 8% in 2024. The 2022 CHIPS and Science Act,[44] which was designed in part to revitalize American basic research, was unable to fill the gaps.[45]

The mismatch between the priority that U.S. policymakers now place on U.S. AI leadership and funding allocations to federal agencies is a disjuncture from AI's First Wave.

That said, one program is particularly important for the U.S. government's role in AI research today: the NSF's National AI Research Institutes.[46] Program Lead James Donlon explained that this relatively new program is central to the U.S. government's AI R&D strategy.[47] Crucially, the Institutes adopt a "use-inspired research framework" that seeks solutions for domain-specific applications where current approaches fall short.[48] Institutes are encouraged to plan for long-term,

The Team NimbRo Rescue semi-autonomus robot uses a power tool to cut through drywall during the Defense Advanced Research Projects Agency (DARPA) Robotics Challenge at the Fairplex June 5, 2015 in Pomona, California. (Chip Somodevilla / Getty Images)

interdisciplinary research projects that contribute to the AI objectives in the National AI R&D Strategic Plan.[49] This includes strengthening and expanding public-private partnerships across government agencies, non-governmental organizations, academia, and industry.[50] Institutes also emphasize complementarity: The federal government's role is to take on "high-risk, high-reward projects" while recognizing that the private sector excels in advanced technology offerings, understanding market trends, and providing access to data and computational resources.[51]

The institutes' mandate thus sidesteps the messiness of Strategic Computing's debate by ensuring their domain-specificity and diffusion while also linking basic research with applications where state-of-the-art techniques do not suffice and emphasizing public-private partnerships and complementarity.

The NAIIO, as well as the National Science and Technology Council's Subcommittee on Machine Learning and Artificial Intelligence,[52] oversee the AI R&D Interagency Working Group (IWG).[53] The IWG, for its part, coordinates and supports long-term investments in AI R&D and applications geared toward U.S. leadership and global competitiveness.[54] This, in turn, includes support for the National AI Research Institutes.[55]

The establishment of the AI Research Institutes was mandated by the National Artificial Intelligence Initiative Act of 2020 ,[56] which calls on the NSF to "lead Federal agencies in providing investments to jump-start … innovations through National AI Research Institutes."[57] Indeed, the institutes are the NSF's flagship program for foundational and use-inspired AI research and the largest research ecosystem funded through partnerships between federal agencies and industry leaders, with $500 billion in total investment across 500 collaborative organizations globally as of 2023[58] and 27 institutes in operation.

Nevertheless, if the disappointments of the First Wave are repeating in the Second, one should be able to trace the fundamental and persistent shortcomings of machine learning through the present day. Indeed, these shortcomings can be identified in fundamental areas including reasoning and planning, abstraction and generalization, factual accuracy, and analytic depth. This undermines AI systems' ability to deliver performance guarantees and provide output of reliability sufficient to justify their critical uses, representing a historical repeat in its own right: In 1984, criticism of Strategic Computing acknowledged the capabilities of AI but cautioned that it "creates a false sense of security" given AI systems' propensity to "act inappropriately in unanticipated situations" owing to a "fundamental limit on their reliability."[59]

Concomitantly, one should be able to detect the U.S. government's role in affirming the perception that intelligence's threshold is being crossed – and one does, particularly in the flow of hardware and export controls therein.

Yet, machine learning's capabilities and limitations are two sides of the same coin; the latter a corollary of the former. Thus, before approaching these limitations, we must understand the Second Wave's deep learning revolution and its intertwining with geopolitics.

The Second Wave's Deep Learning Revolution

The deep learning revolution in part results from the success of a computer vision model[60] known as "AlexNet" in a 2012 image recognition contest.[61] AlexNet outcompeted previous approaches that relied on manually coding features of an image[62] (i.e., symbolic approaches). It instead used the newfound access to vast amounts of data to learn how to discern those features during training. AlexNet marked a break, then, from the First Wave into the Second when "[c] omputation and scale are much more important than human knowledge" in the construction of AI systems.[63] Think of AlexNet as a proof of concept for a realistic alternative to symbolic AI.

During its training, AlexNet performed 4.7 x 10^{17} floating-point operations – this merely refers to the number of times two numbers are added together or multiplied, though this sheer amount of computation was enormous (roughly four-hundred and seventy quadrillion operations!)[64] Past approaches to image recognition did not tend to be as computationally intensive. Indeed, Central processing units (CPUs) were not up to the task of efficiently handling these operations. Thus, AlexNet was trained using two graphics processing units (GPUs), a specialized chip for high-quality image and video processing. An emphasis, then, on data and specialized hardware was present right from the start.

This revolution seeped into strategic reasoning AI. In 2016, DeepMind's AlphaGo, a Go-playing system, defeated international professional player Lee Sedol in four out of five games,[65] exceeding observer expectations. AlphaGo accomplished this with one foot in symbolic AI and the other in machine learning. It linked a problem-solver search algorithm with two deep neural networks.[66] The beauty lies in the interaction between this search algorithm and the neural nets. The search algorithm branches out in a tree-like formation to simulate possible game moves and pathways. The neural networks increase the efficiency of the search algorithm by guiding it towards higher-probability moves (those moves that one network deemed more likely to be made based on the current board state) and then precisely evaluating moves.

The strength of these networks is how they were trained. Two sources of data were involved. The first was data of Go moves made by human expert players. The other was data produced via self-play; that is, data produced by the networks by playing against copies of themselves. AlphaGo's networks were trained on 50 GPUs for one week and three weeks, respectively.[67] During runtime, a "distributed" version of AlphaGo that leverages multiple machines used 1,202 CPUs and 176 GPUs[68] – an indication that not only does the scale of computation matter but also that specialized hardware is demanded.

The self-play technique, formally known as self-play reinforcement learning, took center-stage in 2017 with DeepMind's AlphaGo Zero. It defeated

Facebook accelerates research efforts in Germany on artificial intelligence and machine learning by presenting at the Axel Springer Award in Berlin on Feb. 25, 2016. (Kay Nietfeld / AFP via Getty Images)

AlphaGo, the system that bested Lee, 100-0 – a stunning improvement.[69]

AlphaGo Zero doubles down on deep learning. Specifically, AlphaGo Zero's architecture was simplified to just a single deep neural network. More than this, the network was trained without using examples of human expert moves. As before, the network was trained through self-play, rewarded for wins and punished for losses. Importantly, AlphaGo Zero still possesses the problem-solving search algorithm. The original interactivity between the search algorithm and the neural network in the predecessor AlphaGo was carried into AlphaGo Zero.[70]

AlphaGo Zero's neural network was trained on both GPUs and CPUs and used specialized processors during runtime designed by Google.[71] The emphases on data and specialized hardware persist.

AlphaGo Zero marks a step change from IBM's 1997 chess-playing Deep Blue, which in part relied on internal knowledge related to positions and lines of attack that were hand-coded directly into the system,[72] and received significant input from human grandmasters.[73] Indeed, the AlphaGo research paper explicitly distinguishes Deep Blue's reliance on handcrafted rules from AlphaGo's learning via data.[74]

The momentum for deep learning is most evident in natural language processing (NLP). Historically, NLP's grand challenges relate to the Turing Test,[75] in which a computer, competing against a flesh-and-blood human, convinces a second human it is a real person through anonymized conversation. The successes of the generative pre-trained transformer (GPT) architecture changed attitudes about natural language conversation with machines, with the transformer invented by Google researchers in 2017.[76] GPTs improved with scale through OpenAI's GPT-2[77] and GPT-3,[78] unveiled in 2019 and 2020, respectively. GPT-3's apparent fluency is reminiscent of machines that could pass the Turing Test.[79]

GPT-3's research paper explicitly emphasizes the importance of increasing the size of the model from

GPT-2 to GPT-3, its training dataset size and diversity, and the length of training,[80] reinforcing the association between the scale of computation and specialized hardware with capabilities.[81]

November 2022's ChatGPT-3.5[82] was built on a modified version of GPT-3. OpenAI released GPT-4[83] in March 2023. The model is widely believed to follow the "scaling up" trend, though OpenAI declined to share technical details.[84] Nonetheless, Epoch AI estimates that computing power for AI training increased by almost eight orders of magnitude between AlexNet and GPT-4.[85] The emphasis on specialized hardware persists into 2025, with major companies planning a combined spend of $320 billion on AI and data center build-outs.[86]

Suffusing Machine Learning and Geopolitics

U.S. policymakers have shown increasing interest in AI throughout its Second Wave as the accomplishments under the deep learning revolution accrue and access to hardware becomes intimately linked with progress. There is a distinction, to be sure, in the scope and urgency of U.S. AI policymaking before and after ChatGPT-3.5, though the policy pathway paralleled the industry's apparent growth before this watershed.

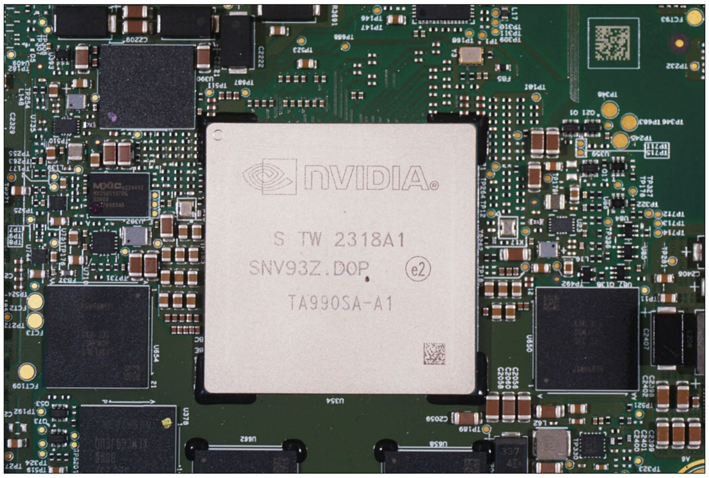
The emphasis that U.S. officials place on restricting the flow of AI-related hardware to China follows the adversaries' great-power competition. The first administration of President Donald Trump oversaw the escalation of a U.S.-China trade dispute that had simmered since the George W. Bush and Barack Obama administrations, intertwining with the Chinese acquisition of sensitive American technologies.[87] U.S. concerns about Chinese access to advanced semiconductors manifested with a pressure campaign on the Dutch government, reported in January 2020,[88] to block sales of chip manufacturing technology to China. The Dutch government decided not to renew the export license for semiconductor equipment maker ASML's extreme ultraviolet lithography (EUV) machine,[89] over which it has supply chain dominance (and depends partly on American technology, giving export controls force.)[90] In May 2020, the first Trump administration amended the Foreign Direct Product Rule (FDPR) to restrict the shipment of semiconductors from global chipmakers to Huawei.[91]

Displaying continuity,[92] in September 2022 the administration of President Joe Biden instructed[93] Nvidia and Advanced Micro Devices (AMD) to cease exporting Nvidia's A100 and H100 chips and AMD's MI250 chips to China – each used in AI development. In October 2022, mere weeks before ChatGPT-3.5 debuted, the Bureau of Industry and Security[94] published an extensive array of export controls designed to restrict Chinese firms from obtaining advanced semiconductors and chipmaking equipment, including a ban on the export of certain chips to China made anywhere in the world with U.S. equipment.[95]

That was pre-ChatGPT. Now, the imperative to gain access to hardware, infrastructure, and energy is more pronounced. [96] In September 2024, BlackRock and Microsoft[97] shared plans to launch a $30 billion private equity fund, dubbed the Global AI Infrastructure Investment Partnership, to build data centers and energy infrastructure to meet AI demand.[98] Abu Dhabi-based MGX, a state-backed AI investment vehicle, is a general partner in the fund alongside Microsoft.[99] Relatedly, in late 2024 Microsoft and Google reached agreements with Constellation Energy and Kairos Power, respectively, to purchase nuclear energy.[100]

As developments unfold, the U.S. has continuously adapted its export controls. The outgoing Biden administration released its "Diffusion" Framework in January 2025.[101] The Framework is comprehensive, dividing the world into three tiers of most to least U.S.-aligned. It also builds on its Data Center Validated End User (VEU) program, allowing companies to apply for National or Universal VEU applications.[102]

Moreover, the U.S. has engaged allies to harmonize restrictions on advanced chips to China. Following the Dutch government's January 2023 restriction on the export of deep ultraviolet lithography machines to China,[103] the Dutch government would, in August 2024, align itself[104] with the U.S. (after some wrangling[105]) by withholding the renewal of ASML's licenses to service and provide spare parts for 1970i and 1980i[106] DUV immersion tools. December 2024's exemption of the Dutch and Japanese, but not states like



An automotive-grade chip developed by NVidia is seen at MWC 2024 in Shanghai, China. (Long Wei / Feature China/ Future Publishing via Getty Images)

South Korea, in its application of the FDPR followed this patchy history.[107]

U.S. officials are highly attuned to the increased demand for advanced chips and infrastructure. OpenAI CEO Sam Altman is trying[108] to persuade officials and investors[109] to pour billions of dollars into AI infrastructure, including financing of new data centers[110] and (at one point[111]) a new chip-building venture,[112] to fuel the large-scale deployment of AI systems. At the White House in September 2024, he pitched the idea that "Infrastructure Is Destiny" and new AI data centers costing $100 billion each should be built – urgently – as a means of reindustrialization.[113]

In November 2024, OpenAI representatives presented a "blueprint for U.S. AI infrastructure" in Washington, D.C, envisioning an infrastructure build-out for AI, complete with state and federal co-created economic zones, a National Transmission Highway Act, and a North American AI Alliance proposal grounded in competition with China.[114] The "Stargate" data center project, jointly announced with Trump in January 2025, might be considered a very partial manifestation of the effort, with government playing a de-regulatory – rather than direct funding – role.[115]

Is Machine Learning the Holy Grail?

A critical mass of American policymakers and officials, then, are locked into the idea that the machine learning paradigm, and more specifically deep learning, is the future of this technology. Retaining American AI leadership – defined by machine learning – thus requires deference to the infrastructural needs of its developmental trajectory; its scaling up. Altman summarizes this sentiment: "In three words: deep learning worked. In 15 words: deep learning worked, got predictably better with scale, and we dedicated increasing resources to it."[116]

This view is seriously problematic, and U.S. policymakers must confront its deficiencies. First, headline-grabbing accomplishments are often more limited than they appear. Second, standards of achievement for AI systems – what counts as a system being "capable" of something – are dramatically lower than in traditional computer science applications. Finally, systems that do merit the descriptor "superhuman" are often more isolated than promoted, not portending future developments that can be seamlessly applied from one domain to another.

What policymakers need today is a view of the machine learning landscape that identifies these shortcomings without dismissing the capabilities this paradigm has achieved (what the architects of Strategic Computing, and their Congressional backers, likewise needed). This requires some level of technical engagement. This is provided below, exploring areas including reasoning and planning, abstraction and generalization, factual accuracy, analytic depth, and intellectual autonomy.

The Misperception of Boundless Innovation

An Arizona State University (ASU) research group led by Subbarao Kambhampati[117] tested the reasoning and planning abilities of large language models (LLMs) from 2022 to 2024, finding that they lag well behind humans: GPT-3 exhibited "dismal performance" when initially tested. [118] A follow-up test found that, while GPT-4 had improved performance over its predecessor by reaching roughly 35% accuracy in a test that requires it to generate plans for stacking blocks ("Blocksworld"),[119] it averages a mere 12% success rate in generating executable plans across domains.[120]

Kambhampati thus likens LLMs' performances to approximate retrieval: LLMs have access to internet-sized datasets, yet unlike a traditional database that faithfully retrieves data exactly as it is stored, LLMs complete an input by reconstructing said data in a probabilistic fashion to generate an output. The ensuing novelty of the output merely looks as though the model is reasoning.[121]

Still, LLMs' purported reasoning abilities often rest on their benchmark scores. Yet, researchers Martha Lewis and Melanie Mitchell highlight the lack of robustness of these scores. They test the analogical reasoning abilities of LLMs – this includes problems that require human and LLM subjects to transfer the abstract structure of one problem to another (e.g., given an original story, participants must judge which of two separate stories are more or equally analogous to the original).[122] When models including GPT-3, GPT-3.5, and GPT-4 are tested on variants of tasks on which LLMs previously performed well – despite their abstract structures remaining the same – LLMs display "brittleness on most of the variation and biases we tested."[123] LLMs' lack of robustness indicates that when LLMs do perform on a par with humans, it is merely because they encountered sufficiently similar problems in their training data, whereas humans appear capable of overcoming their biases through "metacognitive deliberation."[124]

Other problems persist in GPT-based systems. Hallucinations – inaccurate or fictional outputs that LLMs sometimes produce – could be an indefinite problem.[125] Some researchers argue that hallucinations are structural and there is no possibility of ensuring complete accuracy even with access to perfect, up-to-date data.[126] Additionally, even if hallucinations were eliminated, LLMs' responses – particularly in critical applications – still lack sufficient analytical depth.[127]

Would further scaling up – the deep learning revolution's secret ingredient – remedy these flaws? This is unlikely. Research released in April 2024 testing multimodal models – those trained on multiple modalities other than text, like images – finds that the increased performance of the model on a new problem is utterly dependent on how many times the relevant concept appears in its training dataset – and even an exponential increase in training data yields only linear improvements in capabilities.[128] Put simply: More training data may not be enough for the desired capabilities.

Beyond hallucinations, the abstraction and generalization abilities of LLMs are likewise not adequately improved by training on multiple modalities. GPT-4's text-only and multimodal features lack the robust ability to form abstractions relative to humans.[129] On an abstract visual reasoning benchmark, designed with inspiration from human child psychology, multimodal LLMs (including GPT-4V, Claude 3 Opus, Claude 3 Sonnet, and Gemini) give a near-random performance, lagging 40% behind humans.[130]

Finally, testing LLMs on problems related to the pressures of their training environment and its objective – to predict the next word based on the statistical distribution of words in a dataset – find that LLMs' accuracy "can indeed vary substantially" depending on the probability of the example tested.[131] Put simply: LLMs perform better or worse depending on the likelihood of their encountering the type of problem during training, rather than reasoning through them independently.

Unsurprisingly, then, LLM scores on the ARC-AGI-1 Prize[132] – a competition based on the 2019 Abstraction and Reasoning Corpus for Artificial General Intelligence (ARC-AGI) that assessed the capacity for "skill acquisition" and adaptation to a changing environment[133] – are disappointing. On the public, noncompetitive version of ARC-AGI, Claude 3.5 scores 21%, whereas GPT-4o scores 9%.[134]

All this points to fundamental problems in contemporary AI. A paper co-written by Peter Voss, who co-coined the term "AGI,"[135] argues that LLMs are premised on an approach that is fundamentally inconsistent with the original concept of AGI. The focus of the field "shifted from having internal intelligence to utilizing external intelligence (the programmer's intelligence) to solve particular problems."[136] LLMs are woefully unable to autonomously acquire new skills, instead dependent on the instructions, guides, and clues provided by intelligent humans to leverage the resources they possess.

(L–R) Jeff Seibert, Co-founder and CEO of Digits, Kevin Weil, CPO, OpenAI, and Kate Rooney of CNBC speak at the HumanX AI Conference 2025 in Las Vegas, Nevada on March 10, 2025. (Big Event Media / Getty Images for HumanX Conference)

Do OpenAI's 'o1' Models Lay Our Fears to Rest?

To be sure, recent developments ostensibly aim to cure AI's ailments. OpenAI's newest "o1"[137] models are allegedly capable of "reasoning."[138] The company says its o1-preview model performs comparably to doctoral students on benchmark challenges in physics, chemistry, and biology.[139] Both reinforcement learning and "chain-of-thought" (CoT) reasoning are used in o1's design and training.[140] CoT is a technique in which a model is prompted to break down problems into "intermediate natural language reasoning steps that lead to the final output"[141] (i.e., breaking down a problem step-by-step).

OpenAI declined to share architectural details about o1. Plausibly, a modified version of an LLM (e.g., GPT-4o) is pre-trained on data of CoTs; examples of useful reasoning steps expressed via natural language. This model is now capable of predicting the most likely CoT based on the given prompt – think of this as the distribution of CoTs. A reinforcement learning model is then coupled with the modified LLM to hone the distribution of CoTs. Using a specified reward signal (a la AlphaGo Zero), this model generates, selects, and extends a CoT, effectively prompting itself to lengthen the reasoning steps, refining its selection over time.[142]

Whatever the case, when o1 responds to end-users' queries, the response times are unusually lengthy because it is expanding the steps in the "thought process" for improved accuracy. Thus, OpenAI did not move away from the "scaling up" trend but instead applied it to the time during which the model generates outputs.[143]

The reasoning models are both sufficiently different from earlier LLMs to justify a delineation between them and fundamentally deficient in the ways outlined above. The visible throughline is an improvement along some capability measures – say, higher scores on benchmarks – without concomitant improvements in reliability and performance guarantees, factual accuracy, reasoning (names notwithstanding), planning, analytical depth, and so forth.

On a public (non-competitive) version of the ARC-AGI-1 test, o1-mini scored 13% and o1-preview scored 21% (equal to Claude 3.5s, though higher than GPT-4o's 9%).[144] ARC Prize co-founder Mike Knoop explained that the extended CoT prompting does improve the model's ability to adapt to novelty, though o1-preview's parity with Claude was achieved by taking nearly 10 times longer.[145] (On "o3," see below.)

Furthermore, Apple researchers tested 25 state-of-the-art LLMs, including the o1 models, on their logical reasoning capabilities. The researchers did this cleverly: They took a grade-school mathematics benchmark and generated new variants of its mathematical reasoning problems, allowing the researchers to test LLMs through various setups of the questions (much like Lewis' and Mitchell's tests above). For example, one experiment changed the proper nouns (e.g., names) and the numbers of a problem without changing their actual meanings. Other experiments inserted additional clauses into the problems, some relevant and others irrelevant to their required reasoning steps.[146]

On problems where clauses were inserted to increase the difficulty of the problem, all models exhibited performance decreases and variance increases – accuracy diminished, variability ticked up. The rate at which accuracy dropped increased in tandem with the increasing difficulty of the problem. The models' pattern-matching is simply less robust as difficulty increases.[147] When irrelevant, inconsequential clauses were inserted into problems, all models exhibited "catastrophic performance decline,"[148] indicating that models are reliant on pattern-matching the data on which they have been trained.

Interestingly, this research converges on other work in finding that o1-mini and o1-preview exhibit significant improvements over earlier LLMs, yet retain their fundamental shortcomings. The Apple researchers carefully note that o1-preview is not prone to the same type of performance drop and variance on difficulty increases as o1-mini and other models. Yet, both models show a significant performance drop on those problems where irrelevant clauses are inserted into the problems[149] – indicating a lack of genuine logical reasoning. Similarly, the researchers who found that LLMs' accuracy is susceptible to the probability of a given task found that the o1-preview shows "substantial improvement" over previous LLMs, but it continues to exhibit the "same qualitative behavioral patterns that we observed with other LLMs."[150] As other researchers poignantly note, these so-called foundation models "do remain interestingly fragile, especially to unforeseen situations…"[151]

Similarly, the ASU research group tested o1's ability to plan, showing a marked improvement over past LLMs. Testing on three variants of the Blocksworld test, o1-preview performs exceptionally well on the version with complete knowledge of the problems (97.8% accuracy), less well on a version with incomplete knowledge (52.8% accuracy), and a poorer result on an altered, randomized version of the test (37.3% accuracy). These results blow LLMs like Claude 3.5 Sonnet out of the water.[152]

Yet, the retainment of fundamental limitations continues, this time with a First Wave twist. Contrast o1-preview's performance on Blocksworld with a far cheaper, less computationally intensive symbolic planner. This system, Fast Downward,[153] achieves 100% accuracy on all Blocksworld planning tests – perfect scores across the board. The researchers emphasize that Fast Downward accomplishes this in "a fraction of the time, compute, and cost, while providing guarantees that their answers are correct."[154]

That last part is worth our focus. The deep learning revolution is accompanied by a lower standard of achievement for AI systems; they are often claimed to possess a capability, yet they are unable to guarantee the performance that would be expected of said capability. An LLM "can" provide factual, conversation-like text, but it cannot do so reliably; a "reasoning" model like o1-preview "can" plan but it cannot match the performance of a preceding system – in what sense can both Fast Downward and o1-preview "plan?" Computer science applications are traditionally expected to provide "performance guarantees."[155] Deep learning systems often do not.[156]

The Second Wave's mantra is that models like o1 may lag behind systems like Fast Downward in narrow domains, but these are more general models – capable of more than mere planning. Yet, fundamental shortcomings persist as costs of entry rise. New models do not summarily move in a single direction. Nor do models that achieve new capabilities offer performance guarantees one expects from their narrower symbolic predecessors or from a system deserving of the name “artificial general intelligence.”

Note on Confusion Surrounding OpenAI’s “o3”

In December 2024, OpenAI announced its “o3” model.[157] Partnering with ARC-AGI, it is claimed that the model effectively conquered the benchmark with a score of 87.5% using “high-compute” and 75.7% with lower compute.[158] For our purposes, o3 is directionally significant – it likely extends the qualitative trend in o1 of limited adaptation to novelty, but without resolution of fundamental shortcomings.

Public commentary[159] produced confusion about these results. Like o1, o3 was tested on the public ARC-AGI leaderboard. Public leaderboard scores are verified against a semi-private evaluation set to produce a final score.[160] This is the weaker version of ARC-AGI given that some exposure to the data on which the model is tested is assumed to have leaked into its training (thus potentially inflating its score). The claim that it “solved” ARC-AGI-1 is inaccurate absent testing on the private evaluation set.

The high score also excludes compute restrictions, limiting its significance to novelty-adaptation under uncertainty[161] – that it required this compute indicates the system cannot bootstrap its way into new solutions for problems without a helping hand, so to speak. Additionally, OpenAI explicitly trained o3 on the publicly available training dataset[162] – this is a standard practice in machine learning, though inconsistent with the open-ended generalization ARC-AGI is designed to test (doing so effectively undermines the goal of testing a model’s ability to acquire new skills for new problems, as it has trained on sufficiently similar data).

Thus, o3 is directionally significant, but this does not point toward resolving fundamental shortcomings given its training, its excessive compute, and secrecy[163] over its architecture (making a fuller evaluation difficult).[164] External testing will likely indicate the directional significance in o1.

Nevertheless, o3’s design may be moving in the neuro-symbolic direction (ARC founder François Chollet believes it already is neuro-symbolic[165]). Public spasms of euphoria and doom should not distract

Open AI CEO Sam Altman delivers a speech during the “Transforming Business through AI” event in Tokyo, Japan, on Feb. 3, 2025. (Tomohiro Ohsumi / Getty Images)

policymakers from understanding that much more work needs to be done – the capability measures that have marked AI’s progress from AlexNet to o3 are not sufficient for enduring American leadership.

The U.S. Can Lead the Third Wave of AI

A Third Wave of AI development is needed: The strongest contender for this is neuro-symbolic AI. This approach seeks to build on the strengths of the first two waves while mitigating their shortcomings.

Scientific revolutions tend to exhibit a “conservativism”[166] in that they preserve the things worth preserving in earlier paradigms while simultaneously transforming the current understanding. The first two waves produced techniques worth preserving. Indeed, Artur d’Avila Garcez and Luís C. Lamb argue that neuro-symbolic AI should be the Third Wave in which symbolic and neural techniques are coupled to progress on foundational issues.[167]

Precedents exist for neuro-symbolic AI, though they are underplayed.

Researchers explicitly describe DeepMind’s AlphaGeometry – a theorem-proving model – as neuro-symbolic, as it links a rule-based (symbolic)

David Ferris, global head of Cohere, Dan Tadross, head at Scale AI, and Jim Mitre, vice president and director of RAND Global, testify at the Senate Armed Services hearing on artificial intelligence cyber capabilities, on March 25, 2025, in Washington, DC. (Al Drago/Getty Images)

engine with a generative language model (neural).[168] AlphaGeometry 2 and AlphaProof[169] follow this hybrid[170] design. Gary Marcus[171] suggests that DeepMind's protein structure-predicting AlphaFold[172] – of recent Nobel[173] prestige – also possesses a (downplayed) neuro-symbolic structure.

Meta's[174] Cicero, built to play the strategy game (and longstanding AI challenge) Diplomacy,[175] is a beautifully hybrid system, a collection of specialized modules acting within a prespecified and hierarchical structure to handle planning, intent-formation, and communication with other players.[176] Echoing Deep Blue, input from expert human players was substantively integrated into the construction of the system – not a case of mere learning from the data and scaling up the model. Even AlphaGo, Henry Kautz argues, is a "prototypical" example of neuro-symbolic AI in its coupling of a problem-solver search algorithm with a neural network.[177] Despite the description of AlphaGo Zero as starting "tabula rasa"[178] by DeepMind researchers, Marcus correctly points out that the search algorithm was built-in rather than learned from the data.[179]

Finally, the ASU research group put forward a "generate-test" framework in which LLMs are inserted in a loop with a symbolic verifier, allowing the LLM to generate outputs and then improve their generation using the verifier's feedback as it checks their answers. This framework couples the expressiveness of LLMs and their ability to translate problems between formats (i.e., their relative open-endedness) with the domain-specific verifier to guarantee their accuracy (i.e., performance guarantees in critical domains). The generate-test framework improves LLMs' performance and is applicable to the o1 models.[180]

Existing U.S. Government Interest in Neuro-Symbolic AI

By choosing neuro-symbolic AI, policymakers are in good company. DARPA established its Assured Neuro Symbolic Reasoning (ANSR)[181] program in 2022, seeking to "integrate symbolic reasoning with data-driven learning to create robust, assured, and therefore trustworthy systems" and "repair defects in state-of-the-art " machine learning. This follows DARPA's 2018 announcement that it would $2 billion in Third Wave AI systems capable of adapting to new contexts.[182]

The NSF has expressed (comparatively limited) interest in funding neuro-symbolic research. A 2023 program solicitation for National AI Research Institutes detailing the needs of next-generation AI systems lays out three goals: grounding (understanding and robust engagement with concepts and an ability to reason over them), instructibility (effective human control), and alignment (operations consistent with objective, domain-specific truths and human intentions).[183] Neuro-symbolic AI is listed as one possible approach to accomplishing these goals. The reason bears on the lack of responsiveness of data-centric models to these goals without verifiable confidence in future breakthroughs through these techniques.[184] Deep neural networks and the generative models they have spawned cannot guarantee reliability and explainability.[185]

In July 2021, the NSF Division of Information and Intelligent Systems awarded University of South Carolina AI researcher Amit Sheth a $139,999 grant based on a proposal that explicitly invokes the first two waves of AI, arguing that neuro-symbolic AI is the foundation of the Third Wave.[186] The project focuses on the use of "knowledge graphs." Such graphs, as Desta Hagos and Danda Rawat note, represent the relationships between bits of information, thereby serving as a "structured network of interconnected concepts and entities."[187]

An interesting, if subtle, linkage exists between Sheth's NSF-funded work and the AI Research Institutes. In the special issue of "AI Magazine" in which Institute Program Director James Donlon explained their significance, an article coauthored by Sheth and Manas Gaur appears as an issue highlight. The subject: the coupling of generative language models with symbolic techniques (e.g., knowledge-infused ensembles of language models) for critical applications in health care.[188]

Echoes of the U.S. government's role in the foundations of the First Wave reverberate. The goal is to stake out suitable paths forward today without succumbing to earlier perils.

The following message thus drives the recommendations below: American AI leadership is increasingly defined by machine learning. Deference to the infrastructural needs of this technology (and others) has its benefits – including shoring up domestic semiconductor manufacturing capacity[189] and a 19% projected increase in the U.S.'s capture of private-sector investment in wafer fabrication from 2024-2032 thanks to the CHIPS Act[190] – but algorithmic- and architectural-level breakthroughs will be needed to expand American AI leadership; new ideas, not just new chips.

Recommendations for U.S. AI Leadership

Four recommendations reconceive U.S. AI leadership according to this understanding:

1. **The National Artificial Intelligence Initiative Office should direct the AI R&D Interagency Working Group to prioritize neuro-symbolic AI.**

   The federal AI R&D Interagency Working Group's mandate to promote long-term AI investments that conform with U.S. AI leadership should be leveraged to promote neuro-symbolic AI. The NAIIO, together with the Subcommittee on Machine Learning and AI, should therefore direct the IWG to prioritize investments in neuro-symbolic techniques.

   Such investments should be conceived as laying the foundations for U.S. leadership in the Third Wave, targeting deficiencies in AI systems like factual accuracy, reasoning, and planning, abstraction and generalization, and explainability. These investments should simultaneously be seen as pathways to models capable of robustly supporting applications.[191]

2. **The National Science Foundation should establish a national AI research institute for neuro-symbolic AI.**

   Per the National AI Research Institutes' development thus far,[192] a new institute should be established for neuro-symbolic research with an investment worth up to at least $20 million over five years. The purpose of this institute would be to complement existing work in the private sector by bringing together different research traditions while also expanding the reach of basic neuro-symbolic research for socially relevant applications.

   An institute for neuro-symbolic AI should engage in public-private collaboration in earnest, prioritizing those actors willing to collaborate on innovative research in this emerging paradigm. Corporate partners like Meta and Google DeepMind, which are notable for their willingness to invest in neuro-symbolic research across strategic reasoning (Cicero and possibly AlphaGo), mathematics (AlphaGeometry and AlphaProof), and even biological research (AlphaFold), are leading contenders. Equally important are academic partners like Carnegie Mellon, the University of South Carolina, and others. Finally, since the AI Research Institutes emphasize collaboration with international researchers,[193] forming alliances with researchers and organizations within likeminded states is worthwhile.

Importantly, an institute for neuro-symbolic AI should avoid the pitfalls of Strategic Computing and the perils of over-ambitiousness in program and research design. Such an institute should decidedly not aim for AGI, conceived as the hypothetical endpoint of AI. Instead, basic research should be linked to applications in critical domains where current approaches fall short while ensuring its diffusion across the research ecosystem. Fortunately, the AI Research Institutes, pursue the U.S.'s AI objectives in part through complementarity with the private sector (taking high-risk, high-reward projects) and in part through use-inspired research that takes this link seriously.

The matter cannot be settled here, but critical applications in health care – including tasks related to mental health counseling, diagnostics, and clinical guidance, among others – are prime targets for neuro-symbolic research. These should be seriously considered in the establishment of an institute for neuro-symbolic AI.

3. **The U.S. Congress should fulfill the promise of the CHIPS Act by increasing federal agencies' basic research budgets.**

   The budget cuts for basic research funding at agencies including the NSF, NIH, and DoD – contra CHIPS Act expectations – should be reversed. These agencies must have the funds necessary to not only continue to reap the benefits of AI's Second Wave but also invest in foundational research of a sufficiently interdisciplinary nature for its Third Wave.

   There is historical precedent for the U.S. government over-indulging in AI R&D, with Strategic Computing being the archetypal example. The U.S. must take steps to avoid this fate again in a renewed era of great power competition without losing the vibrancy of its federal research ecosystem. The force of these recommendations is that bodies like the NAIIO and the NSF can secure American leadership in the Third Wave by complementing the progress made in the Second; acting as a source of "patient capital"[194] that firms up the foundations of American power[195] by wisely investing the resources it possesses today so that it has that same luxury tomorrow.

4. **The U.S. Congress and Commerce Department should adopt proactive yet targeted export controls on hardware and models in coordination with partners and allies.**

   The U.S. Congress and Commerce Department should ensure that its export controls on hardware or models are aggressively proactive yet targeted, proportional to the actual capabilities of the AI systems they enable or constitute, and implemented in coordination with partners and allies.

   Export controls are effectively a time-buying mechanism;[196] a necessary tool to blunt Chinese firms' efforts to develop AI models on the scale of their American counterparts. By leading the Third Wave, however, the U.S. can achieve two goals simultaneously: curb Chinese companies' advancements in machine learning – effectively restricting them to the Second Wave – while laying the foundations to reap the benefits of neuro-symbolic AI.

   It also, by implication, positions the U.S.'s frontier research to effectively surmount the longer-term diminishing returns of export controls as innovations beyond compute- and data-intensive machine learning unfold.

These recommendations should not be seen as exhaustive. Nor, furthermore, should U.S. policymakers expect the Third Wave to be free of hype cycles. When and if this time comes, it will be incumbent upon U.S. policymakers to be more vigilant in identifying persistent shortcomings in state-of-the-art neuro-symbolic models and begin looking to the future. But the neuro-symbolic train has not yet left the station.

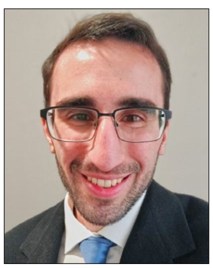
Conclusion

A classic gripe in the machine learning community today is that Marvin Minsky, that pivotal figure in early AI, was so disinterested in the use of ANNs, rather than his favored rule-based systems, that his near-ideological resistance set the field back decades. Imagine, the gripe goes, if neural nets were given their due in the 20 century – LLMs may have been decades old by now!

Putting aside the usual rebuttal to this – that the scaling up required to bring neural networks to their current glory depended on access to hardware that did not exist in Minsky's heyday – the message is clear: Over-indulgence in fundamentally limited symbolic AI harmed the field.

Today, the field risks nurturing a generation of Minskys. This time, machine learning is favored above all else. Their original message, however, remains true: over-indulgence in a fundamentally limited paradigm harms the field. AI is now in the spotlight, a critical technology[197] that promises to be the crown jewel of American technological leadership – such indulgences can no longer be afforded.

America has the resources and the will to lead in AI. It should not squander its opportunity by mistaking machine learning for this technology's endgame.



Vincent J. Carchidi is an analyst focusing on critical and emerging technologies in U.S.-China technology competition and the Middle East. He is a Non-Resident Scholar at the Middle East Institute's Strategic Technologies and Cyber Security Program. He is also a Non-Resident Fellow at the Orion Policy Institute specializing in AI policy. His tech policy work appears in outlets including Defense One, National Interest, The Hill, Trends Research & Advisory, and the Foreign Policy Research Institute. Carchidi maintains a background in cognitive science, applying this research to the trajectories and limitations of frontier AI models. His opinions are his own. You can follow him on LinkedIn, BlueSky, and X.

Endnotes

1 Garcez, A.d. & Lamb, L.C. (2023). Neurosymbolic AI: the 3rd wave. Artificial Intelligence Review, 56(11), 12387-12406. https://doi.org/10.1007/s10462-023-10448-w.

2 Iyengar, R. (2024, August 16). The technocrat. Foreign Policy. https://foreignpolicy.com/2024/08/16/gina-raimondo-us-china-tech-competition-chips-ai/.

3 Hawkins, M. (2024, October 9). U.S. debates who should be able to buy American AI chips. Bloomberg. https://www.bloomberg.com/news/newsletters/2024-10-09/us-debates-whether-to-let-american-chips-fuel-growth-of-global-ai.

4 Hammond, S. (2024). The scramble for AI computing power. American Affairs, VIII(2). https://americanaffairsjournal.org/2024/05/the-scramble-for-ai-computing-power/.

5 Deese, B. and Hansmann, L. (2024, September 12). America needs an energy policy for AI. Heatmap. https://heatmap.news/technology/ai-additionality-framework.

6 Dean, J. (2019). The Deep Learning Revolution and its implications for computer architecture and chip design. ArXiv, 1-17. https://doi.org/10.48550/arXiv.1911.05289.

7 For an accessible overview of the increase in computational demands through the deep learning revolution and how this relates to capabilities, see, Christopher Summerfield. (2023). Natural general intelligence: How understanding the brain can help us build AI. Oxford University Press, 51-60.

8 U.S.-China Economic and Security Review Commission. (2024). 2024 Report to Congress of the U.S.-China Economic and Security Review Commission. U.S. Government Publishing Office. https://www.uscc.gov/sites/default/files/2024-11/2024_Annual_Report_to_Congress.pdf: 27.

9 Zegart, A. (2024, August 20). The crumbling foundations of American strength. Foreign Affairs. https://www.foreignaffairs.com/united-states/crumbling-foundations-american-strength-amy-zegart.

10 Bajraktari, Y. (2024, September 30). The Artificial General Intelligence presidency is coming. Foreign Policy. https://foreignpolicy.com/2024/09/30/artificial-general-intelligence-agi-president/

11 Defense Advanced Research Projects Agency. (1983). Strategic computing: New-Generation computing technology: A strategic plan for its development and application to critical problems in defense (NTIS Issue No.198419). National Technical Reports Library, U.S. Department of Commerce. https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/ADA141982.xhtml.

12 See, Nilsson, N.J. (2009). The Quest for artificial intelligence. Cambridge University Press, ch. 23; Salisbury, E. (2020, May 22). A cautionary tale on ambitious feats of AI: The Strategic computing program. War on the Rocks. https://warontherocks.com/2020/05/cautionary-tale-on-ambitious-feats-of-ai-the-strategic-computing-program/.

13 On this cycle, see, Richbourg, R. (2018, May 10). 'It's either a Panda or a Gibbon': AI Winters and the limits of deep learning. War on the Rocks. https://warontherocks.com/2018/05/its-either-a-panda-or-a-gibbon-ai-winters-and-the-limits-of-deep-learning/.

14 Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R., & Clark, J. (2024). Artificial Intelligence index report 2024. Stanford University Human-Centered Artificial Intelligence. https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf, 244.

15 Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R., & Clark, J. (2024). Artificial Intelligence index report 2024. Stanford University Human-Centered Artificial Intelligence. https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf, 247.

16 Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R., & Clark, J. (2024). Artificial Intelligence index report 2024. Stanford University Human-Centered Artificial Intelligence. https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf, 403.

17 Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J.C., Shoham, Y., Wald, R., & Clark, J. (2024). Artificial Intelligence index report 2024. Stanford University Human-Centered Artificial Intelligence. https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf, 403.

18 Launchbury, J. (2017). A DARPA perspective on artificial intelligence [PowerPoint slides]. https://web.archive.org/web/20161215001731/https://www.darpa.mil/attachments/AIFull.pdf.

19 Defense Advanced Research Projects Agency. (2018). AI next campaign. https://web.archive.org/web/20180908120032/https://www.darpa.mil/work-with-us/ai-next-campaign.

20 National Research Council. (1999). Funding a revolution: Government support for computing research. National Academies Press, 198-205.

21 National Research Council. (1999). Funding a revolution: Government support for computing research. National Academies Press, 199-201.

22 McCarthy, J., Minsky, M.L., Rochester, N., & Shannon, C.E. (1955/2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. AI Magazine, 27(4), 12-14. https://doi.org/10.1609/aimag.v27i4.1904.

23 National Research Council. (1999). Funding a revolution: Government support for computing research. National Academies Press, 202-203.

24 National Research Council. (1999). Funding a revolution: Government support for computing research. National Academies Press, 204.

25 National Research Council. (1999). Funding a revolution: Government support for computing research. National Academies Press, 205.

26 Nilsson, N.J. (2009). The Quest for Artificial Intelligence. Cambridge University Press, 286; see also, National Research Council. (1999). Funding a revolution: Government support for computing research. National Academies Press, 212-213.

27 Symbolic AI is sometimes referred to as "good old-fashioned AI." See, Togelius, J. (2024). Artificial general intelligence. The MIT Press, 84.

28 See, Hagos, D. H. & Rawat, D. B. (2024). Neuro-Symbolic AI for military applications. ArXiv, 2. https://arxiv.org/abs/2408.09224v2; in contrast to the "black box" nature of a deep neural network, a Symbolic system is sometimes called a "white-box system." Feldstein, J., Dilkas, P., Belle, V., & Tsamoura, E. (2024). Mapping the Neuro-Symbolic AI landscape by architectures: A handbook on augmenting deep learning through symbolic reasoning. ArXiv, 3. https://doi.org/10.48550/arXiv.2410.22077.

29 Strickland, E. (2021, September 30). The turbulent past and uncertain future of artificial intelligence. IEEE Spectrum. https://spectrum.ieee.org/history-of-ai.

30 Nilsson, N.J. (2009). The quest for artificial intelligence. Cambridge University Press, 286-287.

31 Feigenbaum, E.A. & McCorduck, P. (1983). The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World. Addison-Wesley.

32 Garvey, C. (2019). Artificial Intelligence and Japan's Fifth Generation: The Information Society, Neoliberalism, and Alternative Modernities. Pacific Historical Review, 88(4), 646. https://doi.org/10.1525/phr.2019.88.4.619.

33 Nilsson, N.J. (2009). The quest for artificial intelligence. Cambridge University Press, 288.

34 National Research Council. (1999). Funding a revolution: Government support for computing research. National Academies Press, 206-207.

35 National Research Council. (1999). Funding a revolution: Government support for computing research. National Academies Press, 207.

36 Salisbury, E. (2020, May 22). A cautionary tale on ambitious feats of AI: The Strategic Computing program. War on the Rocks. https://warontherocks.com/2020/05/cautionary-tale-on-ambitious-feats-of-ai-the-strategic-computing-program/.

37 Salisbury, E. (2020, May 22). A cautionary tale on ambitious feats of AI: The Strategic Computing program. War on the Rocks. https://warontherocks.com/2020/05/cautionary-tale-on-ambitious-feats-of-ai-the-strategic-computing-program/.

38 Rosenblatt, F. (1957). The Perceptron: A perceiving and recognizing automation (Project PARA). Cornell Aeronautical Laboratory, Inc. https://bpb-us-e2.wpmucdn.com/websites.umass.edu/dist/a/27637/files/2016/03/rosenblatt-1957.pdf.

39 See, Haigh, T. (2024, October 8). Between the booms: AI in Winter. Communications of the ACM. https://cacm.acm.org/opinion/between-the-booms-ai-in-winter/#B4.

40 Irwin, J. A. (2024). Artificial worlds and perceptronic objects: The CIA's mid-century automatic target recognition. Grey Room, 97, 17-20. https://doi.org/10.1162/grey_a_00415.

41 The Nobel Prize Organisation. (2024, October 8). The Nobel Prize in Physics 2024. https://www.nobelprize.org/prizes/physics/2024/press-release/

42 National Science Foundation. (2024, October 8). NSF congratulates laureates of the 2024 Nobel Prize in physics. https://new.nsf.gov/news/nsf-congratulates-laureates-2024-nobel-prize-physics.

43 Horowitz, M.C., Allen, G.C., Kania, E.B., & Scharre, P. (2018). Strategic competition in an era of artificial intelligence. Center for a New American Security. https://s3.us-east-1.amazonaws.com/files.cnas.org/hero/documents/CNAS-Strategic-Competition-in-an-Era-of-AI-July-2018_v2.pdf, 6.

44 Badlam, J., Clark, S., Gajendragadkar, S., Kumar, A., O'Rourke, S., & Swartz, D. (2024, May 16). The CHIPS and Science Act: here's what's in it. McKinsey & Company. https://www.mckinsey.com/industries/public-sector/our-insights/the-chips-and-science-act-heres-whats-in-it.

45 Zegart, A. (2024, August 20). The crumbling foundations of American strength. Foreign Affairs. https://www.foreignaffairs.com/united-states/crumbling-foundations-american-strength-amy-zegart.

46 See, National Science Foundation. National Artificial Intelligence Research Institutes. https://new.nsf.gov/funding/opportunities/national-artificial-intelligence-research-institutes

47 Donlon, J.J. (2024). The National Artificial Intelligence Research Institutes program and its significance to a prosperous future. AI Magazine, 45(1), 6. https://doi.org/10.1002/aaai.12153.

48 Donlon, J.J. (2024). The National Artificial Intelligence Research Institutes program and its significance to a prosperous future. AI Magazine, 45(1), 6. https://doi.org/10.1002/aaai.12153.

49 Donlon, J.J. (2024). The National Artificial Intelligence Research Institutes program and its significance to a prosperous future. AI Magazine, 45(1), 6–8. https://doi.org/10.1002/aaai.12153; see also, Select Committee on Artificial Intelligence of the National Science and Technology Council. (2023). National artificial intelligence research and development strategic plan 2023 update. Executive Office of the President of the United States. https://www.nitrd.gov/pubs/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf.

50 Donlon, J.J. (2024). The National Artificial Intelligence Research Institutes program and its significance to a prosperous future. AI Magazine, 45(1), 10. https://doi.org/10.1002/aaai.12153.

51 Donlon, J.J. (2024). The National Artificial Intelligence Research Institutes program and its significance to a prosperous future. AI Magazine, 45(1), 11. https://doi.org/10.1002/aaai.12153.

52 See, Kalil. T. (2016, May 6). Charter of the Subcommittee on Machine Learning and Artificial Intelligence, Committee on Technology, National Science and Technology Council. Executive Office of the President of the United States. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/ai_charter_-_signed_final.pdf.

53 See, Networking and Information Technology Research and Development. Artificial Intelligence R&D Interagency Working Group. https://www.nitrd.gov/coordination-areas/ai/.

54 Subcommittee on the Networking & Information Technology Research & Development and the Subcommittee on Machine Learning & Artificial Intelligence of the National Science & Technology Council. (2024). The Networking & Information Technology R&D Program and the National Artificial Intelligence Initiative Office Supplement to the President's FY2025 Budget. Executive Office of the President of the United States, National Science and Technology Council. https://www.nitrd.gov/pubs/FY2025-NITRD-NAIIO-Supplement.pdf: 20.

55 Subcommittee on Networking & Information Technology Research & Development, Machine Learning & Artificial Intelligence Subcommittee, & National Science & Technology Council. (2023). The Networking & Information Technology R&D Program and the National Artificial Intelligence Initiative Office supplement to the president's FY 2024 budget. Executive Office of the President of the United States. https://www.nitrd.gov/pubs/FY2024-NITRD-NAIIO-Supplement.pdf: 27-28.

56 H.R. 6395 – 1136. National Artificial Intelligence Initiative Act of 2020. https://www.aip.org/sites/default/files/aipcorp/images/fyi/pdf/national-ai-initiative-act-final.pdf: Sec. 5201.

57 Subcommittee on Networking & Information Technology Research & Development, Machine Learning & Artificial Intelligence Subcommittee, & National Science & Technology Council. (2023). The Networking & Information Technology R&D Program and the National Artificial Intelligence Initiative Office supplement to the president's FY 2024 budget. Executive Office of the President of the United States. https://www.nitrd.gov/pubs/FY2024-NITRD-NAIIO-Supplement.pdf, 83.

58 Subcommittee on Networking & Information Technology Research & Development, Machine Learning & Artificial Intelligence Subcommittee, & National Science & Technology Council. (2023). The Networking & Information Technology R&D Program and the National Artificial Intelligence Initiative Office supplement to the president's FY 2024 budget. Executive Office of the President of the United States. https://www.nitrd.gov/pubs/FY2024-NITRD-NAIIO-Supplement.pdf, 83.

59 Ornstein, S. M., Smith, B. C., & Suchman, L. A. (1984). Strategic computing. Bulletin of the Atomic Scientists, 40(10), 12. https://doi.org/10.1080/00963402.1984.11459292.

60 Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2012). ImageNet classification with deep Convolutional Neural Networks. NeurIPS Proceedings, 1-9. https://papers.nips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

61 Dean, J. (2019). The Deep Learning Revolution and its implications for computer architecture and chip design. ArXiv, 1-3. https://doi.org/10.48550/arXiv.1911.05289.

62 Denton, E., Hanna, A., Amironesei, R., Smart, A., & Nicole, H. (2021). On the genealogy of machine learning datasets: A critical history of ImageNet. Big Data & Society, 8(2), 5-6. https://journals.sagepub.com/doi/epdf/10.1177/20539517211035955.

63 Sample, I. (2023, October 28). Race to AI: The origins of artificial intelligence, from Turing to ChatGPT. The Guardian. https://www.theguardian.com/technology/2023/oct/28/artificial-intelligence-origins-turing-to-chatgpt

64 The Economist. (2024, June 30). The race is on to control the global supply chain for AI chips. The Economist. https://www.economist.com/schools-brief/2024/07/30/the-race-is-on-to-control-the-global-supply-chain-for-ai-chips.

65 Moyer, C. (2016, March 28). How Google's AlphaGo beat a Go world champion. The Atlantic. https://www.theatlantic.com/technology/archive/2016/03/the-invisible-opponent/475611/.

66 See generally, Silver, D. et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484–489. https://doi.org/10.1038/nature16961.

67 Silver, D. et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 491. https://doi.org/10.1038/nature16961.

68 Silver, D. et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 487. https://doi.org/10.1038/nature16961.

69 Silver, D. & Hassabis, D. (2017, October 18). AlphaGo Zero: Starting from scratch. Google DeepMind. https://deepmind.google/discover/blog/alphago-zero-starting-from-scratch/.

70 See generally, Silver, D. et al. (2017). Mastering the game of Go without human knowledge. Nature, 550(7676), 354–359. https://doi.org/10.1038/nature24270.

71 Silver, D. et al. (2017). Mastering the game of Go without human knowledge. Nature, 550(7676), 361. https://doi.org/10.1038/nature24270.

72 Goodrich, J. (2021, January 25). How IBM's Deep Blue beat world champion Chess player Garry Kasparov. IEEE Spectrum. https://spectrum.ieee.org/how-ibms-deep-blue-beat-world-champion-chess-player-garry-kasparov.

73 Thompson, C. (2022, February 18). What the history of AI tells us about its future. MIT Technology Review. https://www.technologyreview.com/2022/02/18/1044709/ibm-deep-blue-ai-history/.

74 Silver, D. et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 489. https://doi.org/10.1038/nature16961.

75 Turing, A. M. (1950). Computing machinery and intelligence. Mind, LIX(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433.

76 Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. ArXiv, 1-15. https://doi.org/10.48550/arXiv.1706.03762

77 Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI, 1-24. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

78 Brown, T.B. et al. (2020). Language models are few-shot learners. ArXiv, 1-75. https://doi.org/10.48550/arXiv.2005.14165.

79 This is the overview provided by Summerfield, who notes that the test suffers from being ill-defined and deficient. Yet, GPTs are effectively novel memory algorithms whose information selection is more contextually relevant than past NLP systems. See, Christopher Summerfield. (2023). Natural general intelligence: How understanding the brain can help us build AI. Oxford University Press, 20-24.

80 Brown, T.B. et al. (2020). Language models are few-shot learners. ArXiv, 8-9. https://doi.org/10.48550/arXiv.2005.14165.

81 Garisto, D. (2024, June 4). How cutting-edge computer chips are speeding up the AI revolution. Nature. https://www.nature.com/articles/d41586-024-01544-0.

82 OpenAI. (2022, November 30). Introducing ChatGPT. OpenAI. https://openai.com/index/chatgpt/.

83 OpenAI. (2023, March 13). GPT-4 is OpenAI's most advanced system, producing safer and more useful responses. OpenAI. https://openai.com/index/gpt-4/

84 OpenAI et al. (2023). GPT-4 technical report. ArXiv, 2. https://doi.org/10.48550/arXiv.2303.08774.

85 The Epoch AI Team. (2023, October 23). Announcing Epoch AI's updated parameters, compute, and data trends database. Epoch AI. https://epochai.org/blog/announcing-updated-pcd-database

86 Subin, S. (2025, February 8). Tech megacaps plan to spend more than $300 billion in 2025 as AI race intensifies. CNBC. https://www.cnbc.com/2025/02/08/tech-megacaps-to-spend-more-than-300-billion-in-2025-to-win-in-ai.html.

87 Siripurapu, A. & Berman, N. (2024, May 14). The contentious U.S.-China trade relationship. Council on Foreign Relations. https://www.cfr.org/backgrounder/contentious-us-china-trade-relationship.

88 Alper, A., Sterling, T. & Nellis, S. (2020, January 6). Trump administration pressed Dutch hard to cancel China chip-equipment sale – sources. Reuters. https://www.reuters.com/article/world/uk/trump-administration-pressed-dutch-hard-to-cancel-china-chip-equipment-sale-so-idUSKBN1Z50H4/

89 Clark, D. (2021, July 4). The tech Cold War's 'Most Complicated Machine' that's out of China's reach. The New York Times. https://www.nytimes.com/2021/07/04/technology/tech-cold-war-chips.html

90 The Eurasia Group. (2020). The geopolitics of semiconductors. Eurasia Group. https://www.eurasiagroup.net/files/upload/Geopolitics-Semiconductors.pdf, 8.

91 Shepardson, D., Freifeld, K., & Alper, A. (2020, May 15). U.S. moves to cut Huawei off from global chip suppliers as China eyes retaliation. Reuters. https://www.reuters.com/article/technology/us-moves-to-cut-huawei-off-from-global-chip-suppliers-as-china-eyes-retaliatio-idUSKBN22R1IP/.

92 Soliman, M. & Carchidi, V. (2024, May 16). Technology policy: Convergence and crossroads in Biden vs. Trump 2.0. Foreign Policy Research Institute. https://www.fpri.org/article/2024/05/technology-policy-biden-vs-trump/.

93 Nellis, S. & Lee, J. (2022, September 1). U.S. officials order Nvidia to halt sales of top AI chips to China. Reuters. https://www.reuters.com/technology/nvidia-says-us-has-imposed-new-license-requirement-future-exports-china-2022-08-31/.

94 Bureau of Industry and Security. (2022, October 7). Commerce implements new export controls on advanced computing and semiconductor manufacturing items to the People's Republic of China (PRC) [Press release]. https://www.bis.doc.gov/index.php/documents/about-bis/newsroom/press-releases/3158-2022-10-07-bis-press-release-advanced-computing-and-semiconductor-manufacturing-controls-final/file.

95 Sevastopulo, D. & Hille, K. (2022, October 7). U.S. hits China with seeping tech export controls. Financial Times. https://www.ft.com/content/6825bee4-52a7-4c86-b1aa-31c100708c3e.

96 Hammond, S. (2024). The scramble for AI computing power. American Affairs, VIII(2). https://americanaffairsjournal.org/2024/05/the-scramble-for-ai-computing-power/.

97 Hart, C. & Pitcher, J. (2024, September 17). BlackRock, Microsoft partner on massive new AI infrastructure fund. The Wall Street Journal. https://www.wsj.com/tech/ai/blackrock-global-infrastructure-partners-microsoft-mgx-launch-ai-partnership-1d00e09f.

98 Microsoft Source. (2024, September 17). BlackRock, Global Infrastructure Partners, Microsoft and MGX launch new AI partnership to invest in data centers and supporting power infrastructure. Microsoft. https://news.microsoft.com/2024/09/17/blackrock-global-infrastructure-partners-microsoft-and-mgx-launch-new-ai-partnership-to-invest-in-data-centers-and-supporting-power-infrastructure/.

99 Masters, B., Gara, A., Fontanella-Khan, J., & Morris, S. (2024, September 17). BlackRock and Microsoft plan $30bn fund to invest in AI infrastructure. Financial Times. https://www.ft.com/content/4441114b-a105-439c-949b-1e7f81517deb.

100 Zeitlin, M. (2024, September 20). Microsoft's mega deal is a massive victory for nuclear power. Heatmap. https://heatmap.news/economy/microsoft-three-mile-island-nuclear-constellation; Terrell, M. (2024, October,14). New nuclear clean energy agreement with Kairos Power. Google. https://blog.google/outreach-initiatives/sustainability/google-kairos-power-nuclear-energy-agreement/.

101 Bureau of Industry and Security. (2025, January 13). Biden-Harris administration announces regulatory framework for the Responsible Diffusion of Advanced Artificial Intelligence Technology [Press release]. https://www.bis.gov/press-release/biden-harris-administration-announces-regulatory-framework-responsible-diffusion.

102 Soliman, M. (2025, January 16). Digital borders: The Biden administration's final AI rule. The National Interest. https://nationalinterest.org/blog/techland/digital-borders-biden-administrations-final-ai-rule-214416.

103 Leonard, J. & Koc, C. (2023, January 27). Biden nears win as Japan, Dutch back China chip controls. Bloomberg. https://www.bloomberg.com/news/articles/2023-01-27/japan-netherlands-to-join-us-in-chip-export-controls-on-china.

104 Koc, C. & Leonard, J. (2024, August 29). ASML's China chip business faces new curbs from Netherlands. Bloomberg. https://www.bloomberg.com/news/articles/2024-08-29/the-netherlands-to-put-more-curbs-on-asml-s-china-chip-business

105 Hawkins, M., King, I., Koc, C., & Mochizuki, T. (2024, July 17). U.S. floats idea of tougher trade rules in chip crackdown on China. Bloomberg. https://www.bloomberg.com/news/articles/2024-07-17/us-considers-tougher-trade-rules-against-companies-in-chip-crackdown-on-china.

106 Sterling, T. (2024, September 6). Dutch government retakes export control over two ASML tools from US. Reuters. https://www.reuters.com/technology/dutch-government-retakes-export-control-over-two-asml-tools-us-2024-09-06/.

107 For a more detailed history of US-led export controls and their global interactions, see, Carchidi, V. & Soliman, M. (2024). The role of the Middle East in the US-China race to AI supremacy. Middle East Institute. https://www.mei.edu/publications/role-middle-east-us-china-race-ai-supremacy: 11-27.

108 Hagey, K. & Fitch, A. (2024, February 8). Sam Altman seeks trillions of dollars to reshape business of chips and AI. The Wall Street Journal. https://www.wsj.com/tech/ai/sam-altman-seeks-trillions-of-dollars-to-reshape-business-of-chips-and-ai-89ab3db0.

109 Ghaffary, S. & Ludlow, E. (2024, April 10). OpenAI's Altman pitches global AI coalition on trip to Middle East. Bloomberg. https://www.bloomberg.com/news/articles/2024-04-10/openai-s-altman-pitches-global-ai-coalition-on-trip-to-middle-east

110 Ghaffary, S. & Hawkins, M. (2024, September 3). Altman infrastructure plan aims to spend tens of billions in US. Bloomberg. https://www.bloomberg.com/news/articles/2024-09-03/altman-infrastructure-plan-aims-to-spend-tens-of-billions-in-us

111 Hu, K., Potkin, F., & Nellis, S. (2024, October 30). Exclusive: OpenAI builds first chip with Broadcom and TSMC, scales back foundry ambition. Reuters. https://www.reuters.com/technology/artificial-intelligence/openai-builds-first-chip-with-broadcom-tsmc-scales-back-foundry-ambition-2024-10-29/.

112 Cornish, C. & Murgia, M. (2024, March 15). Abu Dhabi in talks to invest in OpenAI chip venture. Financial Times. https://www.ft.com/content/d018067f-20e7-49eb-83dc-ebb8b1aad1a5.

113 Metz, C. & Mickle, T. (2024, September 25). Behind OpenAI's audacious plan to make A.I. flow like electricity. The New York Times. https://www.nytimes.com/2024/09/25/business/openai-plan-electricity.html.

114 Hayden Field. (2024, November 13). OpenAI to present plans for U.S. AI strategy and an alliance to compete with China. CNBC. https://www.cnbc.com/2024/11/13/openai-to-present-plans-for-us-ai-strategy-and-an-alliance-to-compete-with-china.html; see also, Center for Strategic and International Studies. (2024, November 13). Ensuring U.S. leadership in AI: Industry perspective on data center growth. Center for Strategic and International Studies. https://www.csis.org/analysis/ensuring-us-leadership-ai-industry-perspective-data-center-growth.

115 Goldman, S. (2025, January 22). OpenAI's Stargate may be tech's biggest gamble ever, but here's what's really at stake. Fortune. https://fortune.com/2025/01/22/openai-stargate-ai-sam-altman-donald-trump/.

116 Altman, S. (2024, September 23). The intelligence age. Sam Altman. https://ia.samaltman.com/.

117 See, Kambhampati, S. (2023, September 12). Can LLMs really reason and plan? Communications of the ACM. https://cacm.acm.org/blogcacm/can-llms-really-reason-and-plan/.

118 Valmeekam, K., Olmo, A., Sreedharan, S., & Kambhampati, S. (2022). Large Language Models still can't plan (A benchmark for LLMs on planning and reasoning about change). ArXiv, 2. https://arxiv.org/abs/2206.10498v1.

119 Valmeekam, K., Marquez, M., Sreedharan, S., & Kambhampati, S. (2023). On the planning abilities of Large Language Models: A critical investigation. ArXiv, 5. https://arxiv.org/abs/2305.15771v2.

120 Valmeekam, K., Marquez, M., Sreedharan, S., & Kambhampati, S. (2023). On the planning abilities of Large Language Models: A critical investigation. ArXiv, 2. https://arxiv.org/abs/2305.15771v2.

121 See generally, Kambhampati, S. (2024). Can Large Language Models reason and plan? ArXiv, 1-5. https://doi.org/10.48550/arXiv.2403.04121.

122 Lewis, M. & Mitchell, M. (2024). Evaluating the robustness of analogical reasoning in Large Language Models. ArXiv, 1-31. https://doi.org/10.48550/arXiv.2411.14215. The authors re-test earlier work in addition to their own experiments. For the original work, see, Webb, T., Holyoak, K. J., & Lu, H. (2023). Emergent analogical reasoning in large language models. Nature, 7(9), 1526-1541. https://doi.org/10.1038/s41562-023-01659-w.

123 Lewis, M. & Mitchell, M. (2024). Evaluating the robustness of analogical reasoning in Large Language Models. ArXiv, 22. https://doi.org/10.48550/arXiv.2411.14215.

124 Lewis, M. & Mitchell, M. (2024). Evaluating the robustness of analogical reasoning in Large Language Models. ArXiv, 22. https://doi.org/10.48550/arXiv.2411.14215.

125 Xu, Z., Jain, S., & Kankanhalli, M. (2024). Hallucination is inevitable: An innate limitation of Large Language Models. ArXiv, 1-26. https://doi.org/10.48550/arXiv.2401.11817.

126 Banerjee, S., Agarwal, A., & Singla, S. (2024). LLMs will always hallucinate, and we need to live with this. ArXiv, 1-31. https://doi.org/10.48550/arXiv.2409.05746.

127 See, Gaur, M. & Sheth, A. (2024). Building trustworthy NeuroSymbolic AI systems: Consistency, reliability, explainability, and safety. AI Magazine, 45(1), 142-143. https://doi.org/10.1002/aaai.12149.

128 Udandarao, V., Prabhu, A., Ghost, A., Sharma, Y., Torr, P.H.S., Bibi, A., Albanie, S., & Bethge, M. (2024). No "Zero-Shot" without exponential data: Pretraining concept frequency determines multimodal model performance. ArXiv, 1-41. https://doi.org/10.48550/arXiv.2404.04125.

129 Mitchell, M., Palmarini, A.B., & Moskvichev, A. (2023). Comparing humans, GPT-4, and GPT-4V on abstraction and reasoning tasks. ArXiv, 1-12. https://doi.org/10.48550/arXiv.2311.09247.

130 Jiang, Y., Zhang, J., Sun, K., Sourati, Z., Ahrabian, K., Ma, K., Ilievski, F., & Pujara, J. (2024). MARVEL: Multidimensional abstraction and reasoning through visual evaluation and learning. ArXiv, 1-21. https://doi.org/10.48550/arXiv.2404.13591.

131 McCoy, R.T., Yao, S., Friedman, D., & Griffiths, T.L. (2024). Embers of autoregression show how large language models are shaped by the problem they are trained to solve. PNAS, 121(41), 2. https://doi.org/10.1073/pnas.2322420121.

132 Knoop, M. & Chollet, F. (2024). ARC prize. ARC Prize. https://arcprize.org/

133 Lab42. (2024). ARC: Abstraction & reasoning corpus. Lab42. https://lab42.global/arc/.

134 ARC Prize. (2024). ARC-AGI 2024 high scores. https://arcprize.org/leaderboard.

135 See generally, Goertzel, B. & Pennachin, C. (Eds). (2007). Artificial general intelligence. Springer.

136 Voss, P. & Jovanovic, M. (2024). Why we don't have AGI yet. ArXiv, 1. https://arxiv.org/abs/2308.03598v4.

137 OpenAI. (2024, September 12). Introducing OpenAI o1. https://openai.com/o1/

138 Robison, K. (2024, September 12). OpenAI releases o1, its first model with 'reasoning' abilities. The Verge. https://www.theverge.com/2024/9/12/24242439/openai-o1-model-reasoning-strawberry-chatgpt.

139 OpenAI. (2024, September 12). Introducing OpenAI o1-preview. https://openai.com/index/introducing-openai-o1-preview/

140 OpenAI. (2024). OpenAI o1 System Card. OpenAI, 1-43. https://assets.ctfassets.net/kftzwdyauwt9/67qJD51Aur3eIc96iOfeOP/71551c3d223cd97e591aa89567306912/o1_system_card.pdf.

141 Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2023). Chain-of-Thought prompting elicits reasoning in Large Language Models. ArXiv, 2. https://arxiv.org/abs/2201.11903v6.

142 On this possible architecture, see, Valmeekam, K., Stechly, K., Gundawar, A., & Kambhampati, S. (2024). Planning in strawberry fields: Evaluating and improving the planning and scheduling capabilities of LRM o1. ArXiv, 5, 13-14. https://doi.org/10.48550/arXiv.2410.02162.

143 Kahn, J. (2024, September 13). 9 things you need to know about OpenAI's powerful new AI model o1. Fortune. https://fortune.com/2024/09/13/openai-o1-strawberry-model-9-things-you-need-know/

144 ARC Prize. (2024). ARC-AGI-PUB 2024 high scores. https://arcprize.org/leaderboard#arc-agi-pub.

145 Knoop, M. (2024, September 13). OpenAI o1 results on ARC-AGI-PUB. ARC Prize. https://arcprize.org/blog/openai-o1-results-arc-prize.

146 Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., & Farajtabar, M. (2024). GSM-Symbolic: Understanding the limitations of mathematical reasoning in Large Language Models. ArXiv, 1-5. https://doi.org/10.48550/arXiv.2410.05229.

147 Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., & Farajtabar, M. (2024). GSM-Symbolic: Understanding the limitations of mathematical reasoning in Large Language Models. ArXiv, 8-9. https://doi.org/10.48550/arXiv.2410.05229.

148 Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., & Farajtabar, M. (2024). GSM-Symbolic: Understanding the limitations of mathematical reasoning in Large Language Models. ArXiv, 10. https://doi.org/10.48550/arXiv.2410.05229.

149 Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., & Farajtabar, M. (2024). GSM-Symbolic: Understanding the limitations of mathematical reasoning in Large Language Models. ArXiv, 19. https://doi.org/10.48550/arXiv.2410.05229.

150 McCoy, T.R., Yao, S., Friedman, D., Hardy, MD., & Griffiths, T.L. (2024). When a language model is optimized for reasoning, does it still show embers of autoregression? An analysis of OpenAI o1. ArXiv, 1. https://doi.org/10.48550/arXiv.2410.01792.

151 Lehman, J., Meyerson, E., El-Gaaly, T., Stanley, K.O., & Ziyaee, T. (2025). Evolution and the Knightian blindspot of machine learning. ArXiv, 22. https://doi.org/10.48550/arXiv.2501.13075.

152 Valmeekam, K., Stechly, K., Kambhampati, S. (2024). LLMs still can't plan; can LRMs? A preliminary evaluation of OpenAI's o1 on PlanBench. ArXiv, 1-3. https://doi.org/10.48550/arXiv.2409.13373.

153 See, Helmert, M. (2011). The Fast Downward planning system. ArXiv, 1-56. https://doi.org/10.48550/arXiv.1109.6051.

154 Valmeekam, K., Stechly, K., Kambhampati, S. (2024). LLMs still can't plan; can LRMs? A preliminary evaluation of OpenAI's o1 on PlanBench. ArXiv, 6. https://doi.org/10.48550/arXiv.2409.13373

155 Monroe, D. (2022, October 1). Neurosymbolic AI. Communications of the ACM. https://cacm.acm.org/news/neurosymbolic-ai/.

156 Marcus, G. (2018). Deep learning: A critical appraisal. ArXiv, 14. https://doi.org/10.48550/arXiv.1801.00631

157 Edwards, B. (2024, December 20). OpenAI announces o3 and o3-mini, its next simulated reasoning models. https://arstechnica.com/information-technology/2024/12/openai-announces-o3-and-o3-mini-its-next-simulated-reasoning-models/.

158 Chollet, F. (2024, December 20). OpenAI o3 breakthrough high score on ARC-AGI-PUB. https://arcprize.org/blog/oai-o3-pub-breakthrough.

159 See, e.g., Frazier, K., Rozenshtein, A.Z., & Salib, P.N. (2024, December 23). OpenAI's latest model shows AGI is inevitable. Now what? Lawfare. https://www.lawfaremedia.org/article/openai's-latest-model-shows-agi-is-inevitable.-now-what; see also, Allen, G.C. & Schwartz, A. (2025, January 13). AI diffusion framework – emergency podcast. Center for Strategic and International Studies. https://www.csis.org/analysis/ai-diffusion-framework-emergency-podcast.

160 Knoop, M. (2024, June 27). Introducing the ARC-AGI public leaderboard. https://arcprize.org/blog/introducing-arc-agi-public-leaderboard.

161 Chollet, F. (2019). On the measure of intelligence. ArXiv, 40. https://arxiv.org/abs/1911.01547v2.

162 Misunderstanding arose surrounding a post-hoc note by the ARC-AGI team specifying that OpenAI fine-tuned o3 on 75% of the public training set; ARC-AGI did not test the ARC-untrained model. Given that some data exposure is assumed on the semi-private evaluation, and the ARC-untrained model was not tested, it is not known whether OpenAI fine-tuned its model on leaked evaluation set data, thereby inflating o3's score.

163 In January 2025, information regarding OpenAI's involvement in Epoch AI's FrontierMath benchmark became widely shared, with the former funding the latter and requiring secrecy over the partnership until o3's unveiling while maintaining access to at least some test data (i.e., math problems). See, Meyer, D. (2025, January 21). 'Manipulative and disgraceful': OpenAI's critics seize on math benchmarking scandal. Fortune. https://fortune.com/2025/01/21/eye-on-ai-openai-o3-math-benchmark-frontiermath-epoch-altman-trump-biden/. This case is a cautionary tale and bears some (likely unintentional) resemblance to o3's testing on ARC-AGI.

164 Equally concerning is ARC-AGI's self-contradicting decision not to verify results for a system that scored a baseline performance on the public leaderboard of 83.75% - 85.75%, while nonetheless verifying o3's scores. See, Atreides, K. & Kelley, D. (2024). Solving the Abstraction and Reasoning Corpus for Artificial General Intelligence (ARC-AGI) AI benchmark with ICOM. ResearchGate, 1-30. http://dx.doi.org/10.13140/RG.2.2.32495.34727.

165 Chollet, F. [@fchollet]. (2024, December 30). It's doing CoT search [Post]. X. https://x.com/fchollet/status/1870174117342917041.

166 Weinberg, S. (1974). Reflections of a working scientist. Daedalus, 103(3), 40. https://www.jstor.org/stable/20024218

167 Garcez, A.d. & Lamb, L.C. (2023). Neurosymbolic AI: the 3rd wave. Artificial Intelligence Review, 56(11), 12389-12390. https://doi.org/10.1007/s10462-023-10448-w.

168 Trinh, T.H., Wu, Y., Le, Q.V., He, H. & Luong, T. (2024). Solving Olympiad geometry without human demonstrations. Nature, 625(7995), 476-482. https://doi.org/10.1038/s41586-023-06747-5.

169 Google DeepMind. (2024, July 25). AI achieves silver-medal standard solving International Mathematical Olympiad problems. https://deepmind.google/discover/blog/ai-solves-imo-problems-at-silver-medal-level/.

170 See, Marcus, G. (2024, July 28). AlphaProof, AlphaGeometry, ChatGPT, and why the future of AI is neurosymbolic. Marcus on AI. https://garymarcus.substack.com/p/alphaproof-alphageometry-chatgpt.

171 Marcus, G. (2024, October 9). Two Nobel Prizes for AI, and two paths forward. Marcus on AI. https://garymarcus.substack.com/p/two-nobel-prizes-for-ai-and-two-paths.

172 Jumper, J. et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583-589. https://doi.org/10.1038/s41586-021-03819-2

173 Callaway, E. (2024, October 9). Chemistry Nobel goes to developers of AlphaFold AI that predicts protein structures. Nature. https://www.nature.com/articles/d41586-024-03214-7

174 Meta. (2022, November 22). Cicero. https://ai.meta.com/research/cicero/.

175 Meta Fundamental AI Research Diplomacy Team et al. (2022). Human-level play in the game of Diplomacy by combining language models with strategic reasoning. Science, 378(6624), 1067-1074. https://doi.org/10.1126/science.ade9097.

176 See, Marcus, G. & Davis, E. (2022, November 28). What does Meta AI's Diplomacy-winning Cicero mean for AI? Communications of the ACM. https://cacm.acm.org/blogcacm/what-does-meta-ais-diplomacy-winning-cicero-mean-for-ai/

177 Kautz, H.A. (2022). The third AI summer: AAAI Robert S. Engelmore memorial lecture. AI Magazine, 43(1), 118. https://doi.org/10.1002/aaai.12036

178 Silver, D. et al. (2017). Mastering the game of Go without human knowledge. Nature, 550(7676), 358. https://doi.org/10.1038/nature24270.

179 Marcus, G. (2018). Innateness, AlphaZero, and artificial intelligence. ArXiv, 6-9. https://doi.org/10.48550/arXiv.1801.05667.

180 Valmeekam, K., Stechly, K., Gundawar, A., & Kambhampati, S. (2024). Planning in strawberry fields: Evaluating and improving the planning and scheduling capabilities of LRM o1. ArXiv, 8-10, https://doi.org/10.48550/arXiv.2410.02162; see also, Gundawar, A., Valmeekam, K., Verma, M., & Kambhampati, S. (2024). Robust planning with compound LLM architectures: An LLM-Modulo approach. ArXiv, 1-21. https://doi.org/10.48550/arXiv.2411.14484.

181 Velasquez, A. (2022). Assured Neuro Symbolic Learning and Reasoning (ANSR). DARPA. https://www.darpa.mil/program/assured-neuro-symbolic-learning-and-reasoning.

182 DARPA. (2018). DARPA announces $2 billion campaign to develop next wave of AI technologies. https://www.darpa.mil/news-events/2018-09-07

183 National Science Foundation. (2023, August 1). NSF 23-610: National artificial intelligence (AI) research institutes: program solicitation. National Science Foundation. https://nsf-gov-resources.nsf.gov/solicitations/pubs/2023/nsf23610/nsf23610.pdf?VersionId=DtbAJgpPY438dSJD6rA.55zjuXLlgvnS: 12-13.

184 National Science Foundation. (2023, August 1). NSF 23-610: National artificial intelligence (AI) research institutes: program solicitation. National Science Foundation. https://nsf-gov-resources.nsf.gov/solicitations/pubs/2023/nsf23610/nsf23610.pdf?VersionId=DtbAJgpPY438dSJD6rA.55zjuXLlgvnS: 13.

185 National Science Foundation. (2023, August 1). NSF 23-610: National artificial intelligence (AI) research institutes: program solicitation. National Science Foundation. https://nsf-gov-resources.nsf.gov/solicitations/pubs/2023/nsf23610/nsf23610.pdf?VersionId=DtbAJgpPY438dSJD6rA.55zjuXLlgvnS: 12.

186 National Science Foundation. (2021, June 25). EAGER: advancing neuro-symbolic AI with Deep Knowledge-infused learning. National Science Foundation. https://www.nsf.gov/awardsearch/showAward?AWD_ID=2133842&HistoricalAwards=false.

187 Hagos, D.H. & Rawat, D.B. (2024). Neuro-Symbolic AI for military applications. ArXiv, 4. https://doi.org/10.48550/arXiv.2408.09224.

188 Gaur, M. & Sheth, A. (2024). Building trustworthy NeuroSymbolic AI systems: Consistency, reliability, explainability, and safety. AI Magazine, 45(1), 139-155. https://doi.org/10.1002/aaai.12149.

189 Wu, D. (2024, September 6). TSMC's Arizona trials put plant productivity on par with Taiwan. Bloomberg. https://www.bloomberg.com/news/articles/2024-09-06/tsmc-s-arizona-trials-put-plant-productivity-on-par-with-taiwan?srnd=undefined

190 Varadarajan, R., Koch-Weser, I., Richard, C., Fitzgerald, J., Singh, J., Thornton, M., Casanova, R., & Isaacs, D. (2024). Emerging resilience in the semiconductor supply chain. Boston Consulting Group & Semiconductor Industry Association, https://www.semiconductors.org/wp-content/uploads/2024/05/Report_Emerging-Resilience-in-the-Semiconductor-Supply-Chain.pdf: 4.

191 Sheth, A., Roy, K., & Gaur, M. (2023). Neurosymbolic artificial intelligence (Why, what, and how). IEEE Intelligent Systems, 38(3), 56-62. https://doi.org/10.1109/MIS.2023.3268724.

192 See, Donlon, J.J. (2024). The National Artificial Intelligence Research Institutes program and its significance to a prosperous future. AI Magazine, 45(1), 6. https://doi.org/10.1002/aaai.12153.

193 Donlon, J.J. (2024). The National Artificial Intelligence Research Institutes program and its significance to a prosperous future. AI Magazine, 45(1), 11. https://doi.org/10.1002/aaai.12153.

194 Stewart, N. (2024). Funding for the future: The case for federal R&D spending [White paper]. Special Competitive Studies Project, https://www.scsp.ai/wp-content/uploads/2024/01/RD-White-Paper-2.pdf: 5.

195 See, Zegart, A. (2024, August 20). The crumbling foundations of American strength. Foreign Affairs. https://www.foreignaffairs.com/united-states/crumbling-foundations-american-strength-amy-zegart.

196 Soliman, M. & Carchidi, V. (2024, September 23). Re-Balancing the strategy of tech containment. Foreign Policy Research Institute. https://www.fpri.org/article/2024/09/re-balancing-the-strategy-of-tech-containment/; see also, Carchidi, V & Soliman, M. (2024). The role of the Middle East in the US-China race to AI supremacy. Middle East Institute. https://www.mei.edu/publications/role-middle-east-us-china-race-ai-supremacy: 36-38.

197 Fast Track Action Subcommittee on Critical and Emerging Technologies of the National Science and Technology Council. (2024). Critical and emerging technologies list update. Executive Office of the President of the United States. https://www.whitehouse.gov/wp-content/uploads/2024/02/Critical-and-Emerging-Technologies-List-2024-Update.pdf, 4.