



Artificial General Intelligence: Advancements, Architecture, and Open Challenges

Ms. P. Kowshika, Assistant Professor¹, V. Pratheeksha², J. Divyaa Sri³

¹Assistant Professor, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India

²Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India

³Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India

Abstract -

Artificial General Intelligence (AGI) marks a substantial leap forward within the broader discipline of artificial intelligence. Unlike conventional, task-bound AI systems, AGI is characterised by the capacity to carry out any intellectual undertaking a human being is able to perform, distinguished by adaptability, independent operation, versatility and analytical depth rather than narrow specialisation. This paper offers a detailed examination of AGI, tracing its defining traits, the research directions currently being pursued, and the intricate technical and ethical questions surrounding its continued development. Particular attention is paid to cognitive architectures, neuroscience-informed AI design, and progress in machine learning and deep learning. The discussion further extends to the possible influence of AGI across sectors such as healthcare, education and economic development, together with the corresponding need for robust safety mechanisms, ethical guidelines and governance structures. The paper argues that sustained collaboration among researchers, policymakers and the wider public is essential if AGI is to be designed and deployed for the collective benefit of humanity.

Index Terms — Artificial General Intelligence (AGI), Machine Learning, Cognitive Architectures, Neuroscience-Inspired AI, Deep Learning, Agentic AI.

I. INTRODUCTION

Artificial Intelligence (AI) is undergoing swift transformation, and within this evolving landscape Artificial General Intelligence (AGI) has emerged as a defining and transformative milestone. Narrow AI systems already excel at bounded tasks — image classification, natural language processing, game strategy — yet they fall short of the wide-ranging cognitive flexibility that characterises human intelligence. AGI, by contrast, seeks to build machines that can understand, learn and apply knowledge across

an extensive spectrum of intellectual tasks in a manner comparable to human beings.

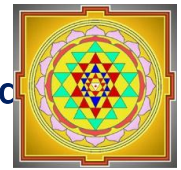
Grasping this distinction matters because AGI carries the potential to reshape entire industries, boost productivity and generate inventive solutions to pressing global problems [1]. Alongside this promise come technical, ethical and societal questions that demand thorough scrutiny and forward planning. Research into AGI remains at an early stage, and considerable advances are still required before anything resembling human-level intelligence is reached [1].

This paper sets out a thorough review of AGI structured to serve both students and researchers. Section II offers a literature review covering definitions, defining characteristics and the range of approaches being pursued toward AGI. Section III lays out an analytical framework for AGI development, illustrated through a conceptual block diagram. Section IV examines the technical, ethical and societal obstacles that stand in the way. Section V presents results and discussion illustrating the predictive capacity of AI-based models across different domains. Section VI closes the paper with a summary of key findings and a look toward future research directions.

II. LITERATURE REVIEW

2.1 Defining AGI and Its Characteristics

Artificial General Intelligence is understood as a system capable of carrying out any cognitive task a human is able to perform. This stands in sharp contrast to narrow AI, which is engineered and trained for a constrained set of functions. Recommendation engines, voice assistants and other specialised systems typify narrow AI: they may outperform humans within their own niche yet break down when asked to operate outside that boundary [1], [8].



The properties most commonly used to define AGI include:

- **Versatility and Adaptability:** the capacity to learn and adjust across varied domains without requiring domain-specific programming, mirroring the flexibility seen in human cognition.
- **Autonomy and Reasoning:** the ability to operate independently, weigh information, reach decisions and resolve problems under shifting conditions without constant human direction.
- **Cognitive Flexibility:** the capacity to absorb, interpret and apply knowledge across many contexts, moving fluidly between unrelated tasks.
- **Continuous and Transfer Learning:** the ability to keep learning from ongoing experience, acquire knowledge independently, and carry insights gained in one domain over into another.
- **Problem-Solving and Creativity:** the capacity for abstract reasoning, inventive problem-solving, and the generation of original scientific or artistic contributions.

Viewed functionally, AI systems are commonly grouped as follows:

- **Reactive Machines:** operate solely on present input with no memory of past events (e.g., IBM's Deep Blue).
- **Limited Memory AI:** draw on historical data to sharpen decision-making (e.g., autonomous vehicles, conversational agents).
- **Theory of Mind AI:** a research frontier in which systems attempt to interpret human emotion, belief and intention.
- **Self-Aware AI:** a speculative future category in which AI would possess consciousness and a sense of self [1], [11].

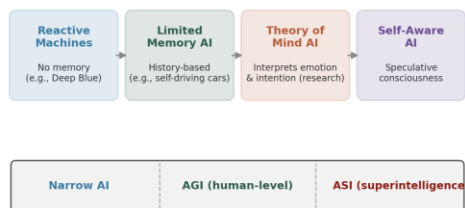


Fig -2: Functional AI Categories Across the Narrow AI → AGI → ASI Spectrum

Whereas AGI targets human-equivalent intelligence, Artificial Superintelligence (ASI) refers to a hypothetical future state in which AI would exceed human capability across every domain.

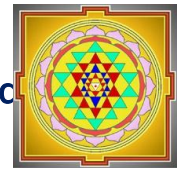
2.2 Approaches to Achieving AGI

A number of distinct research directions are being explored in the attempt to close the gap between narrow and general intelligence:

- **Machine Learning and Deep Learning:** foundational to systems that learn directly from data. Models such as GPT-4 demonstrate impressive linguistic performance, yet remain short of general intelligence, broad applicability and robust generalisation [1], [11].
- **Cognitive Architectures:** initiatives such as OpenCog and the Human Brain Project attempt to replicate human cognitive processes — perception, reasoning, memory — within unified frameworks capable of generalisation and adaptability [13].
- **Neuroscience-Inspired AI:** insights drawn from neuroscientific study of brain function inform the design of more capable AI systems [2], [13].
- **Reinforcement Learning (RL):** trains agents to make decisions through trial and error, deriving effective behaviour from feedback within demanding environments [1], [12].
- **Symbolic AI:** often labelled “Good Old-Fashioned AI” (GOFAI), this approach relies on explicit rules and formal logic, performing well on structured tasks such as legal reasoning while struggling with unstructured data [14], [15].
- **Developmental Robotics:** examines how robots might acquire skills through interaction with their environment, echoing the way human infants learn [11].
- **Hybrid Approaches:** blends symbolic reasoning with the adaptive learning strengths of neural networks and reinforcement learning, helping systems form concepts and justify their actions more clearly [4].

2.3 Current Progress, Benchmarks and Recent Developments

Most researchers regard the route toward AGI as an incremental progression rather than a single decisive breakthrough. OpenAI has proposed a five-tier scale for tracking progress, ranging from Level 1 (basic



conversational AI) to Level 5 (superintelligent AI); the majority of present-day systems remain at Level 1.

Generative AI (GenAI) platforms such as ChatGPT, Gemini and Claude have proven highly capable at producing text, images, code and music. Even so, these systems remain distant from genuine general intelligence, since they typically cannot learn autonomously or grasp meaning and context at a deeper level.

A newer development, Agentic AI, has gained prominence. Rather than responding to single prompts, agentic systems are built to plan and execute multi-step tasks with limited human oversight, particularly within narrow operational domains. Frameworks such as Vectara-agentic and CrewAI support the construction of cooperating AI agents, and agentic AI is widely viewed as a transitional stage bridging today's AI tools and tomorrow's AGI [3], [4], [10].

Since the original review, several further developments are worth noting. Industry and standards bodies have begun treating AGI as a strategic planning category in its own right: the IEEE's 2024 Technology Megatrends report lists AGI alongside digital transformation and sustainability as one of three defining forces shaping the coming decade, while cautioning that the field still lacks a single agreed definition of the "intelligence" that AGI is meant to reach or exceed [16]. Complementing this, broader technical surveys published in 2025 trace how compute scaling, algorithmic advances and increasingly capable foundation models are converging to push the realistic timeline for AGI into a five-to-ten-year horizon, while also flagging unresolved limitations in persistent memory, long-context reasoning and the energy cost of further scaling [17]. Domain-specific reviews have likewise begun examining how AGI-oriented systems, including large language models paired with computer vision, might be adapted to specialised industrial settings such as energy and resource extraction, underscoring that progress toward general intelligence is increasingly being evaluated through practical, sector-specific deployments rather than benchmark scores alone [18].

Key AGI-oriented projects include:

- **OpenCog Hyperon (SingularityNET/TrueAGI):** combines rule-based reasoning with data-driven learning through a shared memory framework called AtomSpace.

- **Zarqa (SingularityNET):** merges deep learning with symbolic reasoning, using knowledge graphs to strengthen language understanding.
- **MuZero (DeepMind Technologies):** attains expert-level performance in complex games without prior knowledge of the rules, relying on search trees combined with learned models.
- **BigMother:** aims to construct a well-aligned superintelligent AGI through a combination of rule-based and learning-based reasoning.
- **OpenAI and Tesla:** both pursue AGI-adjacent goals — OpenAI toward general-purpose intelligence, Tesla toward fully autonomous driving [1], [9], [11].

III. PROPOSED FRAMEWORK — Holistic AGI Architecture

Attaining genuine AGI calls for an integrated design that unites diverse AI techniques, learning strategies and safety mechanisms. The framework depicted in Fig. 1 outlines a modular AGI system engineered for flexibility, continuous learning and responsible operation across domains.

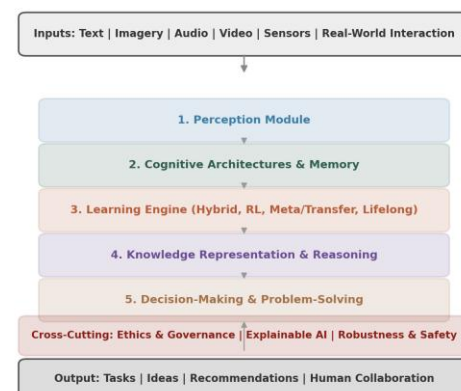


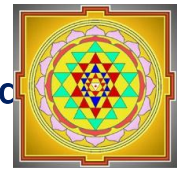
Fig -1: Proposed Holistic AGI Architecture

3.1 Inputs

The system would ingest heterogeneous data drawn from text, imagery, audio, video, physical-world sensors and direct real-world interaction. This diversity of input is essential for constructing a rounded model of the world, much as humans build understanding through multiple senses.

3.2 Core AGI Modules

1. Perception Module: Ingests raw data from varied sources — visual, auditory, textual — and extracts



salient details and context, incorporating advanced visual and audio understanding [1].

2. Cognitive Architectures and Memory: Functions as the system's reasoning core, modelled loosely on the human brain's use of memory, attention and inference. Projects such as OpenCog and the Human Brain Project are working toward systems of this kind [11].

3. Learning Engine: Drives the system's capacity to learn and improve, combining several complementary learning strategies:

- **Hybrid Learning:** pairs pattern-based learning (deep neural networks) with logic-driven reasoning capable of representing cause-and-effect and abstract relationships [8], [12].
- **Reinforcement Learning:** enables learning by trial, feedback and reward, mirroring experiential learning.
- **Meta-Learning and Transfer Learning:** teaches the system to learn more efficiently, acquiring new skills rapidly and applying prior knowledge to unfamiliar situations [14], [15].
- **Continuous/Lifelong Learning:** permits ongoing acquisition of new knowledge without erasing what has already been learned.
- **Self-Upgradation:** allows the system to refine its own methods and capabilities autonomously, guided by what it learns and what tasks demand [13].

4. Knowledge Representation and Reasoning: Organises everything the system learns into structured, usable forms such as knowledge graphs or interconnected concept networks, then applies logical inference, pattern detection and probabilistic estimation to interpret information and resolve problems.

5. Decision-Making and Problem-Solving: Drawing on perception, stored knowledge and contextual understanding, this module determines the most effective course of action, formulates plans and executes tasks — reasoning abstractly, generating novel ideas and adapting to unfamiliar situations.

3.3 Cross-Cutting Concerns

- **Ethical Rules and Governance:** ensures alignment with human values at each stage, encompassing fairness, transparency, accountability and safeguards against harm or misuse [1].

- **Explainable AI (XAI):** ensures the system's decisions remain interpretable to humans, supporting trust and easing review or correction [9].
- **Robustness and Safety:** incorporates strong security measures, redundancy and protective safeguards to shield the system from errors, adversarial attacks or harmful behaviour as it learns and self-improves.

3.4 Output and Interactions

The system carries out tasks, proposes ideas, delivers personalised recommendations and collaborates with human users. It communicates in natural language, operates machinery and interfaces with digital systems, all in a manner consistent with human values.

IV. CHALLENGES

Building and deploying AGI systems entails numerous difficult challenges, which can broadly be sorted into two categories: technical and ethical/societal. Addressing both is essential to realising AGI's benefits while keeping its risks in check.

4.1 Technical Challenges

- **Resource Scalability:** constructing AGI demands vast computing power and data resources, spanning advanced hardware, diverse data modalities and high-speed parallel processing [1].
- **Learning and Model Integration:** AGI must sustain lifelong learning without discarding prior knowledge. Merging multiple learning paradigms across distributed processors poses serious difficulties for scalability and reliability.
- **Control of Self-Upgradation:** ensuring self-improvement remains safe and aligned with human values — particularly in sensitive domains like healthcare or robotics — is exceptionally challenging [8].
- **Perception and Contextual Understanding:** equipping AGI with human-like perceptual abilities is difficult; it must interpret three-dimensional environments, fuse multi-sensor input and grasp context beyond raw data [10].
- **Problem-Solving and Generalisation:** AGI must generalise effectively, applying knowledge from one domain to another and adapting to novel challenges absent from its training data.

4.2 Ethical and Societal Challenges



- **Employment and Economic Impact:** AGI could automate a wide range of jobs, raising concerns over job losses and wage pressure; effective retraining is essential but difficult to implement [1].
- **Privacy and Security:** AGI's capacity to process and generate enormous volumes of data introduces serious risks to personal privacy and data security, alongside potential misuse for cybercrime [5].
- **Accountability and Decision-Making:** assigning responsibility when autonomous outcomes go wrong becomes difficult, especially in healthcare contexts; the "black-box" nature of many models further undermines transparency [6].
- **Control and Alignment:** ensuring AGI behaves safely and remains consistent with human goals is a major challenge, since human values themselves vary and sometimes conflict.
- **Emotional and Social Abilities:** smooth human-AGI interaction requires the system to interpret emotional and social cues, yet emotion is deeply personal and context-dependent.
- **Bias and Fairness:** since AGI learns from historical data, it risks absorbing existing social biases, potentially producing unfair outcomes in hiring, lending or healthcare [7].
- **Governance and Regulation:** robust ethical guidelines and governance frameworks are needed to steer AGI development, including international agreements against misuse [9].
- **Societal Preparedness and Over-Reliance:** readiness for AGI's disruptive effects varies widely across societies, and excessive dependence risks eroding human critical thinking and communication skills.

V. RESULT AND DISCUSSION — AI versus AGI in House Price Prediction

The underlying data illustrates the relationship between house price and floor area (in square feet). Prices generally rise with area, though fluctuations occur due to hidden variables such as location, demand and nearby amenities.

Narrow AI Perspective

- **Pattern Recognition:** models such as regression or decision trees can associate larger homes with higher prices.
- **Task-Specific Prediction:** given a house's size, the model outputs a price estimate and nothing further.
- **Limitations:** factors outside the training data — schools, transport links, market conditions — go unconsidered, and the model must be retrained for a new city.
- **Black-Box Nature:** the model returns a price without explaining why two comparable houses differ substantially in value.

Artificial General Intelligence Perspective

- **Cross-Domain Learning:** AGI would integrate multiple factors simultaneously — economic conditions, safety, infrastructure, culture — to estimate prices.
- **Contextual Adaptation:** it could generate estimates for an entirely new city without retraining.
- **Reasoning and Causality:** AGI could explain its output, e.g., "This smaller property commands a higher price because of its prime location."
- **Transfer Learning:** knowledge gained in real-estate prediction could transfer to related tasks such as rental-trend forecasting or construction-cost estimation.
- **Continuous Learning:** AGI would keep updating itself as market trends, policy and data evolve — something conventional AI cannot do well.

Comparative Insight

Narrow AI behaves like a calculator, identifying patterns in the data it is given (area → price). AGI, in contrast, resembles a real-estate expert — reasoning across domains, adapting to unfamiliar markets, explaining its conclusions and supporting better-informed decisions [1], [10].

VI. CONCLUSION

Artificial General Intelligence remains one of the most compelling frontiers in AI research, holding the potential to transform numerous fields by moving beyond the constraints of today's narrow systems. Its aspiration to unify reasoning, learning and even emotional understanding across domains could yield substantial societal benefits, from advancing healthcare and education to addressing large-scale challenges such as climate change and economic development.



Realising this potential, however, is far from straightforward. Formidable technical obstacles remain — managing large-scale resources, enabling continuous learning, safely governing self-improvement, deepening contextual understanding and achieving general problem-solving. Equally pressing are ethical and societal concerns: job displacement, privacy and security risks, accountability for autonomous decisions, fairness, and ensuring that AGI remains aligned with human values.

This review indicates that meaningful progress depends on sustained collaboration spanning computer science, ethics, policy and the wider public. Guiding AGI's trajectory responsibly will require continued research, transparency in development practice, and firm ethical and governance frameworks. Future work should give particular emphasis to hybrid models, explainable AI (XAI), and comprehensive strategies for managing the broad societal impact of increasingly autonomous systems [1], [8], [9], [11].

REFERENCES

- [1] G. Yenduri, R. Murugan, P. K. R. Maddikunta, S. Bhattacharya, D. Sudheer, and B. B. Savarala, "Artificial General Intelligence: Advancements, Challenges, and Future Directions in AGI Research," *IEEE Access*, vol. 10, pp. 1-25, 2025.
- [2] X. Li, L. Zhao, L. Zhang, et al., "Artificial general intelligence for medical imaging analysis," *IEEE Rev. Biomed. Eng.*, vol. 18, pp. 113-129, 2024.
- [3] W. Saad, O. Hashash, C. Kurisummoottil Thomas, et al., "Artificial general intelligence (AGI)-native wireless systems: A journey beyond 6G," *IEEE Access*, 2024.
- [4] S. K. Jagatheesaperumal, S. Ali, A. Alotaibi, et al., "Generative AI-enhanced neuro-symbolic quantum architectures for secure communications and networking," *IEEE Netw.*, 2025.
- [5] M. Nasr, M. M. Islam, S. Shehata, F. Karray, and Y. Quintana, "Smart healthcare in the age of AI: Recent advances, challenges, and future prospects," *IEEE Access*, vol. 9, pp. 145248-145270, 2021.
- [6] C. M. Fong and M. Usman, "Machine thinking for telehealth—Toward the next generation of healthcare," in *Proc. IEEE SBBS*, Dec. 2020, pp. 1-3.
- [7] L. Chen, P. Chen, and Z. Lin, "Artificial intelligence in education: A review," *IEEE Access*, vol. 8, pp. 75264-75278, 2020.
- [8] B. Wickramasinghe, G. Saha, and K. Roy, "Continual learning: A review of techniques, challenges and future directions," *IEEE Trans. Artif. Intell.*, 2023.
- [9] T. R. McIntosh, T. Susnjak, T. Liu, P. Watters, A. Ng, and M. N. Halgamuge, "A game-theoretic approach to containing artificial general intelligence," *IEEE Trans. Artif. Intell.*, 2024.
- [10] M. Isaac, O. M. Akinola, and B. Hu, "Predicting the trajectory of AI utilizing the Markov model of machine learning," in *2023 IEEE CCAI*, 2023, pp. 30-34.
- [11] V. K. Nyalapelli, M. Gandhi, S. Bhargava, R. Dhanare, and S. Bothe, "Review of progress in artificial general intelligence and human brain inspired cognitive architecture," in *2021 ICCCI*, 2021, pp. 1-13.
- [12] S. Yang and H. Wang, "Sibols: Robust and energy-efficient learning for spike-based machine intelligence in information bottleneck framework," *IEEE Trans. Cogn. Dev. Syst.*, vol. 16, pp. 1664-1676, 2023.
- [13] N. K. Kasabov, "Neuroinformatics, neural networks and neurocomputers for brain-inspired computational intelligence," in *2023 IEEE SACI*, 2023, pp. 13-14.
- [14] S. Yang and B. Chen, "Effective surrogate gradient learning with high-order information bottleneck for spike-based machine intelligence," *IEEE Trans. Neural Netw. Learn. Syst.*, 2023.
- [15] S. Yang and B. Chen, "SNIB: Improving spike-based machine learning using nonlinear information bottleneck," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 53, p. 7852, 2023.
- [16] IEEE Standards Association, "AI Horizon Scanning — White Paper: Part III — Technology Watch," *arXiv:2411.03449*, 2024.
- [17] R. Sarikaya, "Path to Artificial General Intelligence: Past, present, and future," *ScienceDirect*, in press, 2025.
- [18] J. X. Li, T. Zhang, Y. Zhu, and Z. Chen, "Artificial General Intelligence (AGI) for the oil and gas industry: A review," *arXiv:2406.00594*, Jun. 2024.