# The challenges of AI for the next generations

M Yamin[1], M Rihan[2], M Adil[3], Nandan S[4] Mrs.Lolakshi PK[5]

*Students, Department of Information Science and Engineering[1, 2, 3, 4]*
*Faculty, Department of Information Science and Engineering[5]*
*Alva's Institute of Engineering and Technology, Mijar, Mangalore, Karnataka, India*

## Abstract

*Artificial intelligence has advanced significantly and has had a notable impact on various fields such as finance, healthcare, education, and transportation. By using intelligent systems, complex problems can be solved, tedious tasks can be automated, and decision-making can be improved. However, the growing usage of intelligent systems has led to the emergence of several challenges, including security, safety, fairness, robustness, and energy consumption. These challenges have the potential to affect the effectiveness and reliability of intelligent systems and could lead to serious consequences if not addressed. As artificial intelligence moves from narrow to super intelligence, the challenges it poses, including security, will need to be viewed in a different light. Moreover, designing intelligent systems with human-level intelligence is only possible by addressing all the challenges that come with it. Despite this, a comprehensive summary of the challenges of designing artificial intelligence is currently absent in the literature. The purpose of this paper is to address this gap by reviewing these challenges and answering important research questions about their future dynamics and relationships.*

**Keyword:** *Artificial Intelligence, Artificial General Intelligence, Artificial Super Intelligence, Human-Level Intelligence.*

---

# I. **Introduction:**

The field of artificial intelligence (AI) presents both opportunities and challenges, including security, safety, fairness, energy consumption, and ethics. The definitions of different types of AI, such as artificial narrow intelligence (ANI), artificial general intelligence (AGI), artificial superintelligence (ASI), and human-level intelligence (HLI), are continuously evolving. In the healthcare sector, AI is increasingly being used to enhance patient outcomes and optimize clinical workflows. However, challenges such as data quality, ethical considerations, and regulatory concerns must be addressed to fully leverage the benefits of AI while minimizing risks. To fully realize the potential of AI in various fields, including healthcare, it is necessary to have a comprehensive understanding of its challenges and opportunities, as well as its different classes and their evolution.

While many papers focus on the benefits and capabilities of AI, it is important to address challenges such as security, safety, fairness, energy consumption, and ethics to ensure that the potential risks and negative consequences are minimized.

Another challenge posed by the development of AI is its potential impact on employment and the economy. While AI can create new job opportunities and improve productivity, it may also automate certain tasks and replace human workers, leading to job displacement and inequality. It is important to consider the potential economic and societal impacts of AI, and to develop strategies for mitigating any negative effects. This could include retraining programs for displaced workers, policies to promote inclusive economic growth, and regulations to ensure that the benefits of AI are distributed fairly. Additionally, research into the ethical implications of AI and its impact on society should be prioritized to inform policy decisions and ensure that the development of AI aligns with our values and goals as a society.

## 1.1. Preliminaries:

As this paper centers around the obstacles posed by AI, existing studies on AI are positioned through a classification of intelligent systems into three categories, which are outlined below.

ANI: ANI, or Artificial Narrow Intelligence, pertains to intelligent systems designed for particular tasks such as facial recognition or game playing. These agents are programmed to perform specific functions and are not equipped to identify or tackle unfamiliar tasks independently. They lack self-organization capabilities and do not possess self-awareness.

AGI: The notion of AGI, or Artificial General Intelligence, is not universally defined among AI experts. While most researchers use the term to describe agents with intelligence equivalent to humans, there is no consensus on its precise meaning. AGI is often used interchangeably with HLI, or Human-Level Intelligence

ASI: Bostrom proposed three categories of superintelligence, namely, Speed ASI, Collective ASI, and Quality ASI. Speed ASI pertains to agents with faster processing speeds than humans, while Collective ASI refers to decision-making abilities comparable to that of a group of humans. Finally, Quality ASI refers to agents capable of performing tasks that are beyond human capabilities.

## 1.2. Body:

Narrow AI relies heavily on human input, making it susceptible to human error. If a task is set incorrectly, the machine's efforts will be in vain, leading to inaccurate results. Achieving high accuracy rates in AI is challenging, with human performance consistently surpassing even the most advanced AI models. For example, in image classification tasks, humans achieve accuracy rates of over 99%, while AI models require extensive fine-tuning, optimization, and computing power to come close to similar levels of accuracy. Even with the use of pre-existing models, AI struggles to achieve human-level performance. AI-based systems that inherit problems may prove difficult to manage effectively. Furthermore, certain phenomena, such as strange loops, cannot be adequately handled by current mathematical frameworks used by machines.

To utilize learning algorithms in cognitive sciences, the mathematical foundations must first mature. Challenges such as controllability, predictability, security, and safety become increasingly significant during the development of AI, particularly with regards to AGI, HLI, and ASI Super intelligent AI systems pose risks if not designed and used carefully. They could find plans with side effects that go against our interests, such as taking control of physical resources that we would prefer to use differently. If these systems acquire more power than we do, it could lead to a shortage of resources for humans. If a super-intelligent AI system does not respect our values, it could cause a global catastrophe or even human extinction. The superintelligence control problem is the challenge of understanding and managing these risks. While achieving super intelligent systems in the next few decades is unlikely, studying the problem of Superintelligence control is worthwhile.

AI has revolutionized software development by providing a powerful way to gain trust through verification. However, in critical areas such as medical and military services, lack of code verification in AI-based systems may be unacceptable. Since these systems often act as 'black boxes' and lack transparency in their predictions and decisions, ensuring that they are validated is critical. To address this issue, there is an increasing focus on developing algorithms to visualize, explain, and interpret AI models. For example, efforts are underway to evaluate the use of learning algorithms to detect breath sounds in children with lung disease in a real clinical setting. Given the importance of validation, we can expect a large number of publications covering both theory and applications in the near future. Nonetheless, verifying the self-assembly capabilities of HLI-based agents is not a trivial task.

Reducing the complexity of AI-based systems has become a challenging task as they now incorporate numerous learning models for intra-modular recognition and decision-making processes. This complexity is amplified by the expansive parameter space resulting from multiplying the parameters of the system's internal components. Addressing these challenges, researchers have proposed various approaches, such as AutoML and parameter optimization methods. However, dealing with complexity in Human-Level Intelligent (HLI) agents is particularly

intricate due to the involvement of a substantial number of agents and the complexity of human intelligence, which remains largely unexplored.

## II.CONCLUSION:

This paper focuses on analyzing the challenges of AI in the context of HLI. It's important to note that the challenges vary in popularity, with some such as security and fairness being more widely discussed than others like energy and complexity. To provide a comprehensive understanding of each challenge, a brief description was given, and the connections and evolutions between them were studied. However, some well-known challenges like the curse of dimensionality were not discussed due to their popularity. As we continue to evolve from ANI to ASI, new dimensions of challenges may emerge. While some challenges like the potential monopoly of corporations and job loss were not covered, they could be addressed in future research. The paper emphasizes that a multidisciplinary approach is necessary to address these challenges, involving not only computer science but also social sciences, ethics, and policy-making.

One prominent challenge not discussed in the paper is the issue of bias in AI systems, which can lead to discrimination and unjust treatment for certain groups. To tackle this challenge, we need to increase diversity in the development and training of AI systems, along with bias-testing and validation methods.

Another challenge is the potential for malicious use of AI systems, such as cyberattacks or targeted disinformation campaigns. As AI systems advance, they become more sophisticated in carrying out such attacks, making it harder to detect and prevent them. Thus, we require a collective effort from the AI community to develop cybersecurity measures and ethical guidelines.

Lastly, the challenge of accountability is critical in the development and deployment of AI systems. With AI making significant real-world decisions, we need to establish accountability frameworks ensuring transparency, explain ability, and responsibility. This includes addressing ethical implications and liability issues. Addressing these challenges requires a multidisciplinary approach that encompasses ethics, social sciences, and policymaking, in addition to computer science.

## REFERENCES:

1. "The Future of Employment: How Susceptible are Jobs to Computerisation?" by Carl Benedikt Frey and Michael A. Osborne
2. "Superintelligence: Paths, Dangers, Strategies" by Nick Bostrom
3. "Life 3.0: Being Human in the Age of Artificial Intelligence" by Max Tegmark
4. "AI Superpowers: China, Silicon Valley, and the New World Order" by Kai-Fu Lee
5. "The Ethics of Artificial Intelligence" by Nick Bostrom and Eliezer Yudkowsky