DOI:10.1145/3664804

Even as we celebrate AI as a technology that will have far-reaching benefits for humanity, trust and alignment remain disconcertingly unaddressed.

BY VASANT DHAR

The Paradigm Shifts in Artificial Intelligence

ARTIFICIAL INTELLIGENCE (AI) captured the world's attention in 2023 with the emergence of pre-trained models such as GPT, on which the conversational AI system ChatGPT is based. For the first time, we can converse with an entity, however imperfectly, about anything, as we do with other humans. This new capability, provided by pre-trained models, has created a paradigm shift in AI, transforming it from an application to a general-purpose technology that is configurable to specific uses. Whereas historically an AI model was trained to do one thing well, it is now usable for a variety of tasks, such as general conversations; assistance; decision making; and the generation of documents, code, and video—for which it was not explicitly trained. The scientific history of AI provides a backdrop for evaluating and discussing the capabilities and limitations of this new technology and the challenges that lie ahead.

Economics Nobel Laureate Herbert Simon, one of the pioneers of AI who coined the term "artificial intelligence", described it as a "science of the artificial."[43] In contrast to the natural sciences, which describe the world as it exists, a science of the artificial is driven by a goal: creating machine intelligence. According to Simon, this made AI a science of design and engineering. Just five decades later, pre-trained models have greatly expanded the design aspirations of AI, from crafting high-performing systems in narrowly specified applications, to becoming general-purpose and without boundaries, applicable to anything involving intelligence.

The evolution of AI can be understood through Kuhn's theory of scientific progress[22] in terms of "paradigm shifts." While philosophers of science have debated Kuhn's multiple uses of the term "paradigm" in his original

» key insights

- **Kuhn's framework of scientific progress helps us understand how artificial intelligence has gone through several paradigm shifts during its scientific history, and how the representation and use of knowledge has widened in scope and scale.**
- **The current general intelligence paradigm has breached a major barrier where machines can learn through self-supervision, transforming AI from an application to a general-purpose technology like electricity that will enable the rapid development of transformative applications across the economy.**
- **In the current AI paradigm, lack of trust is the key barrier, which must be overcome for the widespread adoption of AI.**

ILLUSTRATION BY PETER CROWTHER ASSOCIATES

Figure. The history of artificial intelligence.

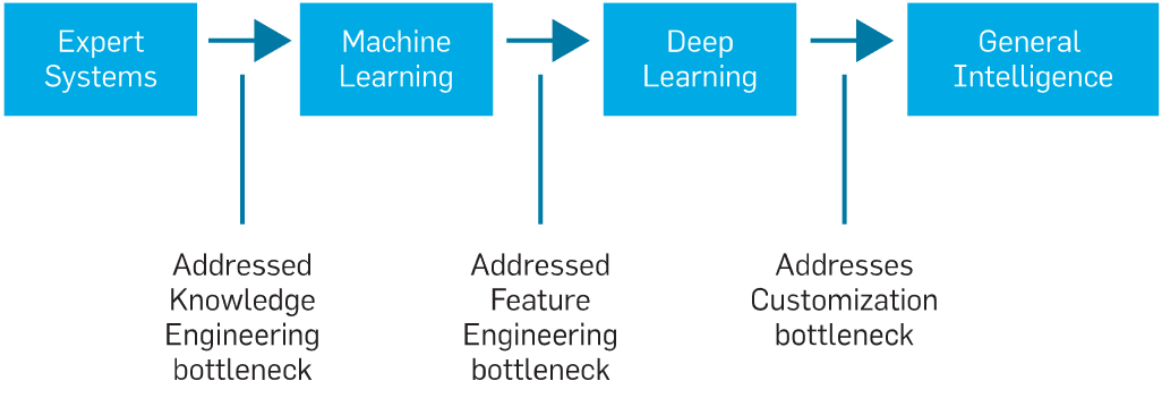

Table 1. The paradigm shifts in AI.

	Data	Exemplar	Scope	Curation
Expert Systems	Human	Rules	Follows	High
Machine Learning	+ Databases	Rules/networks	+ Discovers relationships	Medium
Deep Learning	+ Sensory	Deep neural networks	+ Senses relationships	Low
General Intelligence	+ Everything	Pre-trained deep neural networks	+ Understands the world	Minimal

NOVEMBER 2