

# Trustee Insights

## EMERGING ISSUES



## The Impact of Artificial Intelligence in Health Care

As Artificial Intelligence grows, trustees should consider its pros and cons

BY STEVEN M. BERKOWITZ, M.D.

**H**ardly a day goes by without at least one story in the news or on social media about artificial intelligence (AI). We are bombarded with topics both inspiring and terrifying. AI has become a buzz word, an overused and inappropriately cited cliché. But there is no doubt that this technology will radically change how we do things in the future, particularly in health care. In the face of all this publicity and hype, how can the trustee separate the wheat from the chaff? What is the role of the trustee

in providing leadership and strategic guidance for the organization?

The trustee is faced with a double challenge — understanding the implications of AI in one's own field as well as the health care profession. Fortunately, the AI fundamentals apply to both. We begin with a working definition and practical model for AI. Familiarity with these basics puts the trustee in a better position to make strategic decisions amidst the plethora of potential products and applications that the organization will encounter, and help answer the following questions: What do we mean by AI in the first place? How do we conceptualize and define it? Can AI develop a

The following is Part I of a two-part series on artificial intelligence (AI) for the trustee. The first article will discuss the definition of AI, provide a practical working model and cover general concepts and controversies that the Trustee will encounter. Part II, coming in September's Trustee Insights, will be devoted exclusively to health care and will discuss the need for an organizational partnership with AI, the importance of incorporating AI into the organization's strategic planning process and specific health care applications -- both clinical and operational.

life of its own? Is this technology a blessing or a curse? Can it save humanity or destroy humanity? The list goes on.

### A working definition of the AI Process

A working definition would be:

"Artificial intelligence is the process whereby humans program a computer with a set of problem-solving algorithms and supply it with a database upon which to apply these algorithms. The computer then provides a solution to the problem, simulating human intelligence."

This definition appears simple, straightforward and safe. But there is a problem. In the actual application of AI, two phenomena have been empirically observed that raise concerns:

The first is known as emergent properties. Although specific algorithms or “rules” are initially programmed into the computer, the computer may randomly start to deviate from these “rules” and produce results or conclusions that are entirely unexpected and unpredictable. One could say the computer could develop a “mind of its own” which the original human programmers cannot predict or control. The ability to develop emergent properties allows the “machine” to modify the human controls and supervision. This introduces a scary randomness or uncertainty into the above definition — a well-defined and confined process of human-machine engagement. There is current debate about the etiology of this phenomenon and how frequently it occurs, but we must concede and manage this operational reality.

Secondly, AI can randomly give a confident response that is nonsensical or not justified by its training data or source content. These events are referred to as hallucinations. In other words, it can make a false statement randomly and unpredictably. Put another way, the computer on occasion “will make things up,” begging the question as to whether AI can be trusted at all. Again, it is not clear how often these hallucinations occur and whether they can be prevented, but like the emergent properties, they must be considered in the final AI output and managed.

Therefore, as a practical matter, we must amend our initial definition of AI to include the observed concepts of emergent properties and hallucinations. This makes the impact of AI much more precarious and frankly scarier, as humans have

the potential to lose control of the technology. Here is an amended and unsettling definition:

### Realistic Definition

**A process whereby humans give a computer access to algorithms and data. The computer applies these algorithms to the data and generates a solution imitating human intelligence ...**

**... BUT randomly, the computer may provide an unjustified or false solution**

### A practical descriptive model for artificial intelligence

Complicated AI programs and applications can be better understood and analyzed by breaking down the process into three components. Each component can evolve separately and influence the progress of the others. The ultimate AI output is a function of all three operating together as a system:

1. Hardware
2. Software
3. Connectivity

#### 1. Hardware — the exponential development of computer processing power

There has been unprecedented progress in the past sixty years in terms of fulfilling the mantra: faster/smaller/cheaper. Moore’s Law, coined in 1965 by Intel’s Gordon Moore, stated that the number of transistors that can be put on an integrated circuit will double about

every two years. This doubling process has continued unabated to this day. Processing speed and efficiency are essential in the evolution of AI. It has been the rate-limiting step in the development of more advanced capabilities such as language processing and complex machine learning. It has only been in the last five or six years that we have had the processing capability to have such entities as GPT and deep learning applications. True advanced intelligence is now operationally and commercially feasible.

One can only imagine what will be possible as computer power continues to double. Quantum computers, for example, which rely upon sub-atomic particles to power the processing, are currently in the research stage, and promise to continue the cycle of “faster/smaller/cheaper.” Emerging photon computers will use light instead of electricity. Together, they will allow AI to operate millions of times faster and require less energy. Operationally, Moore’s Law is alive and well for the near future.

#### 2. Software — the evolution of machine learning and complex reasoning:

The second component involves the application of machine learning and complex reasoning through ever more complex and sophisticated software. Supervised machine learning, where models, trained with labeled data sets and pre-set algorithms, can progress to unsupervised machine learning, using new, real-world data. This allows the AI to evolve and move beyond the confines of initial data sets and algorithms. As algorithms and systems

become more sophisticated, more complex logic that simulates or even exceeds human capacity is destined to become a reality.

It is now possible to program the most basic form of reasoning, deductive reasoning (going from the general to the specific) as well as the more complicated inductive reasoning (going from the specific to the general). The question is now whether the computer can “machine learn” more advanced forms of human reasoning such as creativity, emotions and even empathy. Most believe this is possible and, in some instances, this has already been achieved, creating a blurring of what a machine can do and what a human can do. At this rate, it is only a matter of time before the machine will exceed human intelligence. This new learning can bring enormous benefits, but it also can amplify the negative effects of false outcomes, emergent properties and hallucinations.

**3. Connectivity- the computer→ the human brain→ neural networks→ robotics.**

The third component of the model is interoperability. Each individual computer has the capacity to interface and connect

to other computers and datasets forming large networks. These neural networks allow access to enormous amounts of data from diverse sources such as the internet or the “cloud.” Brain-computer interfaces (BCIs) enable the computer to be directly connected to the human brain, allowing the brain and computer to seamlessly work together. Similarly, the computer and brain can now connect directly to robotic machines. Thus, the boundaries between the computer, the brain, robotics and networks such as the internet are intermeshed.

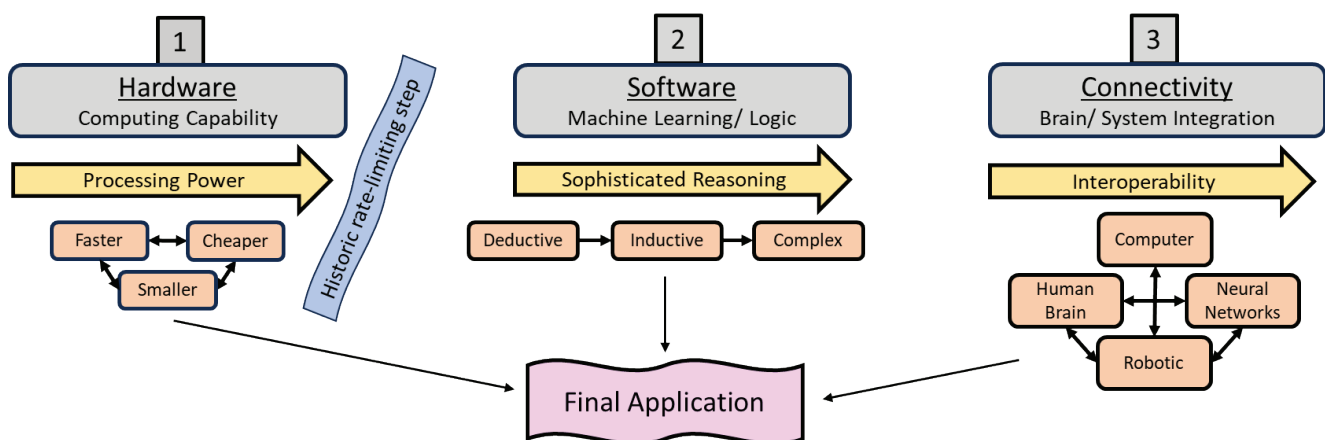
Understanding how these three components – hardware, software and connectivity – work synergistically helps to explain why AI technology is proceeding so fast and so efficiently. This model can be used to objectively evaluate present and future AI applications in a consistent manner. The diagram below summarizes the model.

**Key concepts and controversies with artificial intelligence**

Armed with the fundamental definition and a practical model for AI, let us now address some key concepts

and controversies that trustees will encounter both in their respective professions and as a health care board member.

**Generative pretrained transformer (GPT)** — GPT is an AI-powered chatbot that simulates a human and can generate responses when asked questions. No AI innovation has taken the market by storm in recent years as much as this. Within two months of OpenAI’s release of ChatGPT, it had over one hundred million users. It can compose essays, code computer programs, summarize complex topics and analyze pictures. It has performed very impressively on SAT tests, bar exams and even the Multistate Professional Responsibility legal ethics exam. Additionally, when asked medical questions and compared to physician responses, a multidisciplinary medical team rated its responses higher than physicians on both quality and empathy. The possibilities of GPT applications seem endless. Vastly more powerful updates are on the horizon. Multiple vendors are now entering this space. GPT will be embedded in many processes in all industries. Its potential in health care is overwhelming.



**Deep fakes** — Due to the success of GPT and increasingly sophisticated software, it is now possible to generate images, videos, or audio recordings that are indistinguishable from the originals. This has huge implications in entertainment, music and politics. What is real and what is fake? If someone makes a deep fake of you, who has ownership of that piece? Deep fakes generate controversies in regulation and copyright laws. It remains to be seen where this will land, but it is an area of legitimate concern. Vendors offer the ability to separate real versus AI generated material. Meanwhile, the “bad guys” continue to produce more sophisticated ways to evade detection.

**Inherent biases of AI** — When AI provides an intelligent solution to a problem, could there be intrinsic bias in that solution or is it truly objective? There are several opportunities to introduce bias into the AI process. First, the initial algorithms and training data may have a bias from the programmer either unintentionally, or even deliberately. Most people agree that there is bias in large data sets as well as the information obtained from the internet. Finally, as was discussed, the emergent properties and hallucinations could affect the credibility of the output. A recent article gave Chat-GPT the Political Compass quiz, and it came out significantly on the left and libertarian side. It is fair to assume that any AI output could contain biases from

numerous etiologies, and specific results should always be assessed for this possibility.

**AI and consciousness** — As AI systems become more sophisticated and more intelligent, at what point could it be considered “conscious” or self-aware? For example, GPT responses, as mentioned above, were rated to be more empathic than humans. Does the computer actually “feel” empathic or is it just responding in an empathically programmed fashion? Most researchers feel that the computer presently has not achieved all aspects of the definition of consciousness. Having said that, given the rapidly expanding technology, the potential of crossing over that barrier into full self-awareness and consciousness must be considered. If this can be accomplished, it would be a significant step toward the concept of singularity.

**The concept of singularity** — In technology, singularity describes a hypothetical future where technological growth becomes out of control and irreversible. In the context of artificial intelligence, the singularity would occur when the technology is vastly more intelligent than humans. AI could in theory take over the world. Is this media hype, or is it our fate? No one doubts that the growth of AI continues to be exponential. There is considerable controversy in this area, but by extrapolating present growth curves, some feel that the actual point of singularity could occur as soon as five

to seven years. One of the most primal instincts of a “living” organism is the need to survive. If the computer perceives a human as a threat, would it then feel compelled to destroy that human? Presently, this is the fodder of science fiction novels and movies. However, many respected AI researchers have expressed concern.

It is my hope that most of these topics generate more questions than answers. Collectively, they fuel the hype and paranoia of AI exceeding human intelligence and taking control. This must be weighed against the compelling benefits to humanity that this technology would bring.

It makes sense that 100% of the population should be excited about the tremendous potential of AI. It is equally prudent that 100% of the population should be concerned. It emphasizes the challenge for the trustee to weigh both the pros and the cons of this technology as they direct strategy and provide oversight of the organization.

Artificial intelligence gives us a great deal to think about. And the technology has only begun.

---

**Steve M. Berkowitz, M.D.,** ([steve@smbhealthconsulting.com](mailto:steve@smbhealthconsulting.com)), is, founder and president of SMB Health Consulting based in Scottsdale, Ariz.

---

*Please note that the views of authors do not always reflect the views of the AHA.*