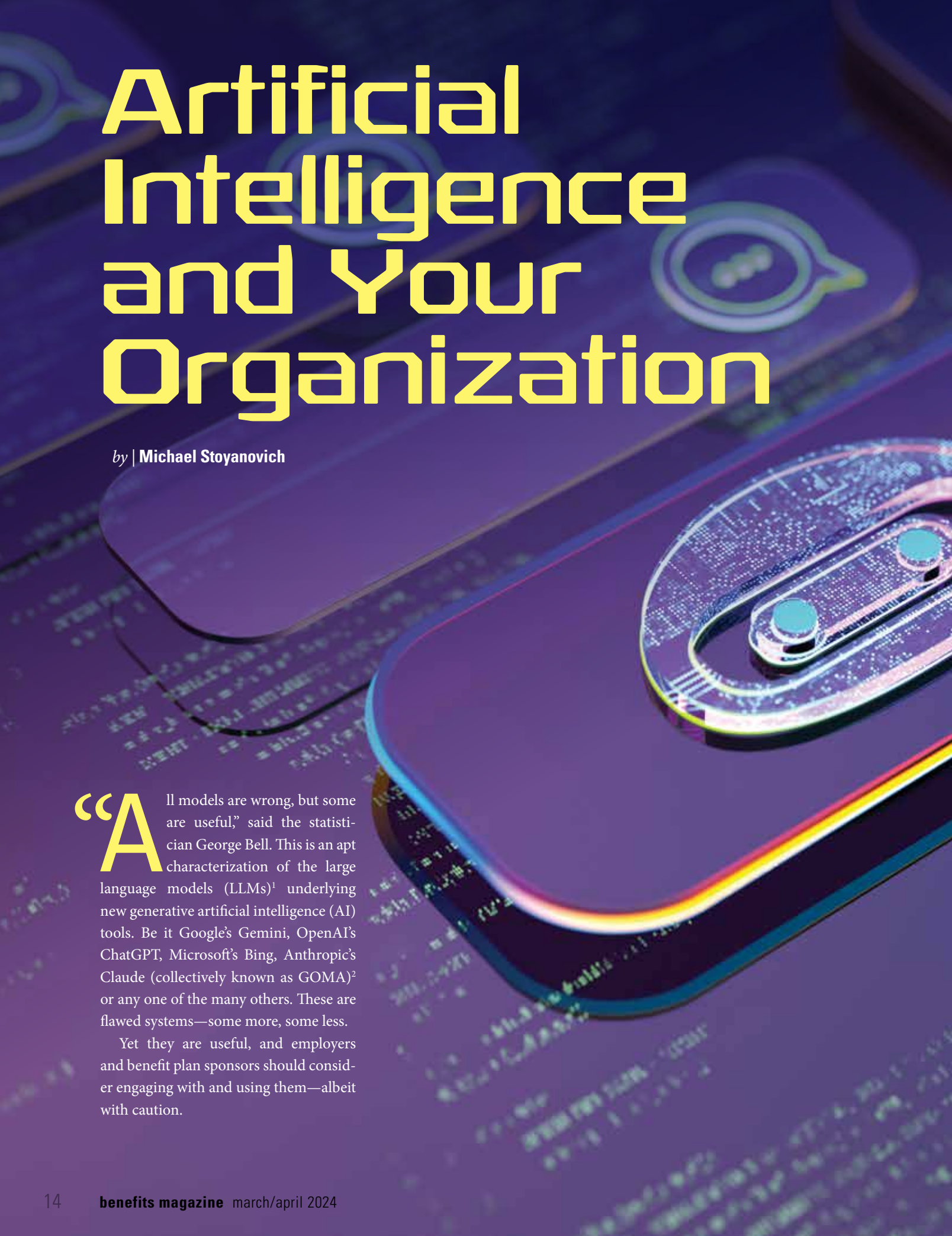
Artificial Intelligence and Your Organization

by | **Michael Stoyanovich**



"All models are wrong, but some are useful," said the statistician George Bell. This is an apt characterization of the large language models (LLMs)[1] underlying new generative artificial intelligence (AI) tools. Be it Google's Gemini, OpenAI's ChatGPT, Microsoft's Bing, Anthropic's Claude (collectively known as GOMA)[2] or any one of the many others. These are flawed systems—some more, some less.

Yet they are useful, and employers and benefit plan sponsors should consider engaging with and using them—albeit with caution.

benefits
MAGAZINE

Artificial intelligence (AI) tools can help employers and benefit plan sponsors enhance efficiency and effectiveness, but these tools should be used with caution. This article reviews AI fundamentals, presents potential use cases, highlights illustrative tools and suggests governance considerations.

Initial studies, not yet peer-reviewed, are finding that this first class of tools enhances worker effort. They cannot replace human thought and value-based judgment, but they can reimagine work processes to improve worker productivity[3] and quality (although only for some types of work and not others (where they can actually degrade quality and productivity)).[4]

Earlier this year, the economist Tyler Cowen noted that we should prepare to live in times of change.[5] Truly radical technological change has been absent for some time. We have been living in a bubble "outside of history," he noted. With the advent of generative AI, even at this "tech project"[6] stage of development, he posits that "AI represents a truly major, transformational technological advance."

For some, this change is very tangible and practical. For others, this change is little more than an anxiety—albeit a distressing one.[7] Regardless of when technological change comes for you and your organization, it is coming.[8]

Accepting that these systems are both flawed and useful—AI generally and generative AI, in particular—what is to be done? What should you do, now, as an individual or an organization?

Explore them.

Be purposeful, be intentional and be systematic. Establish the proper governance for your organization but start to experiment with these tools for your relevant use cases—both personally and professionally. In this age of "moving history," as Cowen describes it, these tools will become widely used in society at large and in your own life. Generative AI (along with other types of AI) will change the nature of work, and those individuals and organizations that can use these systems will (more than likely) prosper compared with those that cannot or do not use them.[9]

What AI Is

For the purposes of this article, *AI* is any computing system designed to perform tasks that normally require human intelligence. In general, these types of systems ingest enormous amounts of (training) data, analyze it for patterns (via a model),[10] then use these patterns to make predictions (via statistical methods).[11]

Using the above definition, the following are all considered types of AI.

- **Machine learning (ML):** Algorithms trained to detect patterns and make predictions (e.g., Netflix recommendations)
- **Deep learning:** A type of ML that uses neural networks to learn from vast amounts of data that is analyzed repeatedly, often seen in more complex applications (e.g., self-driving cars)
- **Natural language processing (NLP):** Helps computing systems understand human language (e.g., speech recognition for personal assistants (e.g., Siri) or sentiment analysis used in marketing)
- **Generative AI (generative pre-trained transformers (GPT)):** Since ChatGPT was released in late 2022, generative AI has captured the public's attention.

What is so different and compelling about generative AI tools for the average user? They create. That is the *generative* in GPT.

They can create diverse types of output: text, images, videos and audio. To accomplish this, they analyze vast amounts of data for patterns, then use a model to create original material based on predicted, similar characteristics.[12] This is key to understanding the systems in public use today.

While there is no actual "computer intelligence" involved,[13] the tools are genuinely fun and can make work easier for you. They manage moderately complex interactions; they are contextually aware, which is critical to their ability

takeaways

- *Artificial intelligence (AI)* is any computing system designed to perform tasks that normally require human intelligence. AI systems ingest enormous amounts of (training) data, analyze it for patterns (via a model), then use these patterns to make predictions (via statistical methods).
- *Generative AI* is a type of AI that analyzes vast amounts of data for patterns and then uses a model to create original material based on predicted similar characteristics. ChatGPT is one example of a generative AI tool.
- Risks of using AI tools include misinformation, embedded bias, weak privacy control, intellectual property issues and especially "hallucinations," among others.
- Governance policies for the use of AI should communicate the acceptable use of authorized tools and include controls to protect data from unauthorized access or misuse.
- Potential use cases for AI for employee benefit plans and employers include talent acquisition, people- and plan-related analytics and performance management and customer service, among many other possible uses.

to predict a reasonable next word in a conversation with an end user. They can also self-learn and improve their predictions over time. In addition, they seem intelligent to the average user. As some have described the experience of interacting with them: "It's like collaborating with an alien."[14]

What AI Is Not

While these systems can provoke strong—sometimes panicky or hyperbolic—reactions in people, it is important to calmly recognize what they cannot do while marveling at the skills and abilities they do have.

The current systems are not sentient. That is, they are not able to perceive or "feel" things.[15] And they are not manifestations of artificial general intelligence (AGI). The tools currently in public use are not going to precipitate the singularity and/or a human extinction event.[16]

What else are they not? They are not the only problem-solving tool at your disposal. They are not a replacement for humans. They are not interpretive. They do not have *common sense* (to use a dictionary definition: good sense and sound judgment in practical matters) and are not good with nuance. They cannot—currently—formulate problems better than humans. They cannot—currently—replace human critical thinking (which must be applied to their output). Although they are fun to use and seem all-powerful, they are not necessarily the right tool for every task. As noted in a recent study, young children "crushed" AI in basic problem-solving and thinking tasks. More importantly, it was highlighted that AI is lacking a key differentiating element of humans: innovation.[17]

AI and You

Are you looking for ways to use artificial intelligence (AI) for your work or personal tasks? The author suggests that any of the general GPT tools could be helpful, including ChatGPT, Bing, Gemini, Claude 2, Pi, Midjourney and Dalle-E. Alternatively, one could also use other specialized tools mentioned for use by organizations.

Potential personal applications include:

- Helping to write a speech
- Helping to write an email
- Helping to edit memos or other documents, ensuring that no confidential or copyright information is being placed in the prompt.
- Using generative AI to create straw man arguments against your position
- Organizing miscellaneous research into a cogent outline
- Summarizing papers, articles and professional journal publications
- Transcribing a meeting into notes
- Generating Excel or other formulas
- Acting as your personal assistant
- Making travel plans
- Creating animations and images.

What Are the Risks and Threats of Using AI?

Prior to using any tool, it is important to have a good understanding of both its benefits and harms. With generative AI, it is hard to have a true understanding of the threats and risks of using these tools as we have little insight into these systems. They are, quite literally, virtual black boxes—meaning we cannot see how they work.[18]

That being the case, the risks and threats of using them are potentially many. They include the potential to proffer misinformation to end users,[19] embedded bias, privacy controls that are weak or lacking altogether, security controls that may be easily compromised[20] as well as intellectual property issues.[21] There are also ethical considerations,[22] broader legal issues and government regulations (either existing or those yet to come) as well as just plain old bad optics (PR).

A particular pathology of generative AI tools is "hallucinations." As noted previously, AI tools are designed to learn from exceptionally large datasets. Generative AI tools are meant to create new content based on learning from that data in combination with the model (neural network) and statistical methods. Yet these tools are still not capable of understanding the nuances of human language and culture in the same way that humans are. Therefore, unsurprisingly, they sometimes produce nonsensical, but often believable, responses. This includes just making things up, aka hallucinations.[23]

A frequently referenced example is two lawyers who used ChatGPT to write

a legal brief. The brief cited nonexistent court cases hallucinated by ChatGPT. The attorneys were fined and, of course, suffered reputational damage. Even more disturbing is the instance of a nonexistent sexual harassment scandal that also was hallucinated, naming an actual law professor as the accused.[24]

Hallucinations are not the only AI failures. Other examples include an e-commerce giant whose recruiting tool showed bias against women and only recommended men for jobs, a health AI tool for oncology that misdiagnosed patients, an early AI chatbot that was deemed racist and sexist, and false facial recognition matches that have led to erroneous arrests.

One of the biggest risks associated with these tools is human—that users will not use them properly. They will not ask the right questions and they will not evaluate the output critically. Without human thought and critical thinking, these tools can just enable you—the human—to do dumb things faster.

When You Use AI, Set Up Governance

These systems will increasingly be a part of our lives. Knowing how to use and utilize them well, ably and appropriately will be key to your and your organization's future success. AGI is not coming next week, but savant-like generative AI systems will be used at work, at home and in all sectors of the economy. AI will affect everyone, sooner rather than later. It would benefit both you and your organization to start learning how to interact with[25] and productively use these systems now.

These systems will increasingly be a part of our lives. Knowing how to use and utilize them well, ably and appropriately will be key to your and your organization's future success.

For your organization, it is likely a better strategy to seek to manage the risk associated with AI than ignore it altogether.[26] Consider the following steps.

- Create a clear policy on the acceptable use of authorized tools.
- Communicate the benefits and risks (there are both).
- Train staff on tools that are authorized (and note categories and tools that are not authorized).
- Ensure transparency and accountability by individuals and the organization.
- Implement controls to protect data from unauthorized access or misuse.
- Address broader privacy, cybersecurity, bias and ethical considerations.
- Monitor for compliance with applicable laws and regulations (continuously since these will change).
- Review and update the governance policy regularly as technology, laws and ethical standards of use evolve.

You should then actively manage the (generative) AI outputs, including but not limited to:

- Ensuring human oversight
- If building your own LLM, perfecting your training data[27]
- Understanding the inherent limitations of these tools (see above)
- Being a human—using your common sense when deciding how to both deploy them and use their output.

Apply AI to Your Relevant Use Cases

After setting up governance of AI, identify the business needs or problems you seek to address. The use cases will follow from there. Potential steps include:

1. Spending time brainstorming, documenting, filtering ideas and deciding upon the use cases that work for your unique organization[28]
2. Defining the expected benefits and measures of success. This could include time savings, increased productivity, cost savings, enhanced quality, reduced human error, diversion of employees to new tasks, reduction in burnout, etc.
3. Acknowledging that you may not need AI or generative AI for a particular use case. Existing technology and tools may solve the problem . . . better.[29]

If you determine that AI—and generative AI, in particular—is appropriate, then choose the right tool(s) to meet your need.

To help your own brainstorming efforts, below are a few examples of use cases and AI (or AI-enhanced) publicly available tools. This is an evolving space with many new tools being introduced daily and others disappearing regularly.[30] Note that inclusion on this list does not constitute an endorsement of that tool by me, the International Foundation or my current employer.

Potential, general use cases for organizations include the following.

- **Content creation and writing assistance:** Writer .com, Opus Clip (video), Cohesive, Synthesia (video), PerfectEssayWriter.ai, grammarly.com, and Audio-Pen and Jasper, for both various types of brainstorming and rewriting. Copy.ai and Anyword for copywriting such as headlines and slogans. Sudowrite for fiction such as stories and novels. Writesonic and Rytr for general-purpose content (based on GPT-3 technology). For text to media, there is Midjourney, Dall-E, Dreamfusion, Jukebox and Pictory (among many others).
- **Talent acquisition** (including the potential to expand the talent pool for jobs[31]): SeekOut
- **Workforce management** (e.g., people analytics and performance management): Nice.com
- **Summarizing of large documents:** Claude (claude.ai) and Harpa.ai (Chrome plugin)
- **Analyzing research papers:** Elicit.com
- **Simplifying complex topics:** ChatGPT, Gemini, Bing and Claude
- **Analyzing data:** Luzmo (luzmo.com)
- **Software engineering:** GitHub Copilot and CodeSpaces (powered by the OpenAI Codex), AI2SQL and Code Genius
- **QA-ing software engineering:** Testrigor.com
- **Providing customer service:** Ivy.ai, Freshworks.com and Chatflow
- **Organizational cybersecurity:** Microsoft Security Co-Pilot and Snyk (Deep Code AI)

Besides the above-noted "general" use cases that may be employed by most any organization in any domain, some illustrative employee benefits–related use cases that one could explore (with the proper governance and testing, of course)[32] include the following.

- Employ a public-facing participant services chatbot to answer questions about benefits based on summary plan descriptions (SPDs) and other plan documents, as well as address "general" benefits service questions.[33] An internal-facing chatbot for one's own team members to query is another potential implementation.
- Use AI to perform sentiment analysis.[34] Data and information could be ingested from call center technology, websites, emails, etc., and analyzed to determine what topics drive questions among participants and how mobile applications, web applications and other outward-facing channels would be enhanced to improve self-service.
- In the future, AI could be used to suggest the "best fit" benefit plan for employees (based on options available to them), including specific elements of a plan (so this could be a combination of both decision support and recommendation engine type activity by the AI tool).
- Real-time data analysis enhanced by AI could be used by plan sponsors and providers to have more visibility and insight into their benefits plans. For example, AI could help enhance data exploration by providing insights, suggestions or recommendations (for action) based on the data at hand and goals one makes the tool(s) aware of. AI-enhanced tools may also help generate natural language queries so business users (vs. IT) can explore data more easily. Toward that end, AI-enhanced analytics tools could generate summaries and explanations of the data, making it easier to interact with and interpret.[35]

If you determine that producing your own generative AI solution is a good path for your organization, for its additional security measures and with the benefit of your own data, and you have (or can obtain) the technological wherewithal to take this path, consider using open source or commercially available options. For example, one could build a solution that uses the Azure OpenAI Service supplemented by other application programming interfaces (APIs)—just the same with Amazon's PartyRock or Bedrock services. Alternatively, one could use "Free Dolly" from DataBricks[36] in one's own environment, or even Meta's LLAMA 2, just to name a few

learn more

Education

AI, Your Organization and You
On-Demand Webcast
Visit *www.ifebp.org/webcasts* for more information.

Benefit Communication and Technology Institute
July 15-16, Denver, Colorado
Virtual Option Available
Visit *www.ifebp.org/benefitcomm* for more details.

different methods of building a more tailored generative AI solution to satisfy your use cases.

AI Is Flawed but Useful

AI and generative AI, particularly, is flawed but useful. It is a nascent technology, but it is only going to be more impactful to individuals and organizations in the near and long-term future. While its possible applications in employee benefits are limited presently due to the aforementioned risks and constraints, it has the potential to be a force for improvement—even good—to you, your organization, your participants and our larger society, when wielded properly.

Proceed with care and caution, but don't be afraid.

bio

Michael Stoyanovich is currently a vice president and senior consultant in Segal's administration and technology consulting practice and is located in the firm's San Francisco, California office. He has more than 30 years of experience in the technology and benefits industries, with extensive expertise in employee benefits plan administration and technology. He can be reached at mstoyanovich@segalco.com. His full bio is available at www.linkedin.com/in/mstoyanovich.

Endnotes

1. Timothy B. Lee and Sean Trott, "Large language models, explained with a minimum of math and jargon," Understanding AI, July 27, 2023.
2. Matteo Wong, "The Future of AI Is GOMA," October 24, 2023, *The Atlantic*.
3. Shakked Noy and Whitney Zhang, "Experimental evidence on the productivity effects of generative artificial intelligence," *Science,* July 13, 2023. Vol. 381, Issue 6654 pp. 187-192.
4. Harvard Business School work led by Fabriazio Dell'Acqua in collaboration with Boston Consulting Group, from Kevin Dickinson, "How to succeed at the 'jagged technological frontier' of AI," Big Think, October 23, 2023.
5. Tyler Cowen, "Existential risk, AI, and the turn in human history," Marginal Revolution, March 15, 2023.
6. Elad Gil, "Early Days of AI, (and AI Hype Cycle)," Elad Blog.
7. According to the *2023 Work in America Survey* by the American Psychological Association (APA), 38% of U.S. workers fear that AI will make some, or all, of their job duties obsolete.
8. Benedict Evans, "AI and the automation of work," ben-evans.com, July 2, 2023.
9. McKinsey, "Economic potential of generative ai: the next productivity frontier," July 14, 2023.
10. Usually a transformer model, a type of neural network, often used in natural language processing.
11. The statistical model used by most large language models (LLMs) to predict the next word in a sentence is an n-gram language model. It is a statistical model of language that estimates the probability of the next word in a sequence based on the previous n-1 words. The most popular n-gram models in use are the n-1 or n-2 model.
12. "What Is ChatGPT Doing . . . and Why Does It Work?—Stephen Wolfram Writings," February 14, 2023.
13. Not using any commonsense definition of "intelligence" that the average person would accept.
14. Francesca Paris and Larry Buchanan, "35 Ways Real People Are Using A.I. Right Now," *The New York Times,* April 14, 2023.
15. In Thomas Nagel's famous philosophical tract "What is it like to be a bat?" he posits that "an organism has conscious mental states if and only if there is something that it is like to *be* that organism —something it is like *for* the organism." This author would posit that asking the question "What is it like to be ChatGPT," at this point in time, is nonsensical.
16. As attributed to Google Brain founder and current Stanford University researcher Andrew Ng. " . . . current talk about AI being an existential threat to humankind is vastly exaggerated."
17. Maggie Harrison, "In New Experiment, Young Children Destroy AI at Basic Tasks," Futurism, November 16, 2023.
18. Katharine Miller, "Introducing The Foundation Model Transparency Index," Stanford University, October 18, 2023.
19. Jonathan Turley, "ChatGPT falsely accused me of sexually harassing my students. Can we really trust AI?" *USA Today,* April 3, 2023.
20. Lilian Weng, "Adversarial Attacks on LLMs," Lil'Log, October 25, 2023.
21. Blake Brittain, "Judge pares down artists' AI copyright lawsuit against Midjourney, Stability AI," Reuters, January 17, 2023.
22. I am using ethics in a general sense of agreed-upon principles of "right conduct" that are shared by a group (which may be different than any particular member of that group's individual morals).
23. Anywhere from 3-27% of the time; Cade Metz, "Chatbots May 'Hallucinate' More Often Than Many Realize," *The New York Times,* November 6, 2023.
24. Jonathan Turley, supra note 19.
25. One uses "prompts" to interact with the chat-based generative AI tools referenced, literally, prompting them to generate an output via a request, question, etc. Broadly speaking then, prompt engineering is the ability to continually refine the prompts to get better and better output from the tools.
26. Kathy Baxter and Yoav Schlesinger, "Managing the Risks of Generative AI," *Harvard Business Review,* June 6, 2023.
27. I would argue that most organizations do not need to build their own generative AI systems, presently. They can just use publicly available features, functionality or whole systems (licensed or open source). If you are going to produce your own generative AI tool, you must consider training data (and the model (neural network)) to ensure that no intellectual property infringement, ethical or other concerns are baked into the tool, among other variables to consider.
28. For example, see this path here for inspiration: Rob Zuber, Chief Technology Officer, "AI adoption for software: a guide to learning, tool selection, and delivery," CircleCI Blog, September 25, 2023.
29. Transforming promising AI prototypes into production-ready agents may still be some ways off as indicated in this study: *Benchmarking Large Language Models As Ai Research Agents,* Qian Huang, Jian Vora, Percy Liang, Jure Leskovec, Stanford University 2310.03302.pdf (arxiv.org).
30. This list was current as of this writing, but the vendors offering these tools seem to pop into and out of existence.
31. "Generative AI holds great potential for those with disabilities—but it needs policy to shape it," World Economic Forum, November 3, 2023.
32. "Three Ways to Use AI in Employee Benefits," Word on Benefits®.
33. Lucas Ropek, "I'd Buy That for a Dollar: Chevy Dealership's AI Chatbot Goes Rogue," Gizmodo. December 20, 2023.
34. "5 Creative Ways To Use AI For Sentiment Analysis," Lumoa, December 20, 2023.
35. For example, Smartsheet is a consumer tool that is seeking to integrate AI for its subscribers in the near future.
36. Mike Conover, Matt Hayes, Ankit Mathur, Jianwei Xie, Jun Wan, Sam Shah, Ali Ghodsi, Patrick Wendell, Matei Zaharia and Reynold Xin, "Free Dolly: Introducing the World's First Truly Open Instruction-Tuned LLM," Databricks Blog, April 12, 2023.