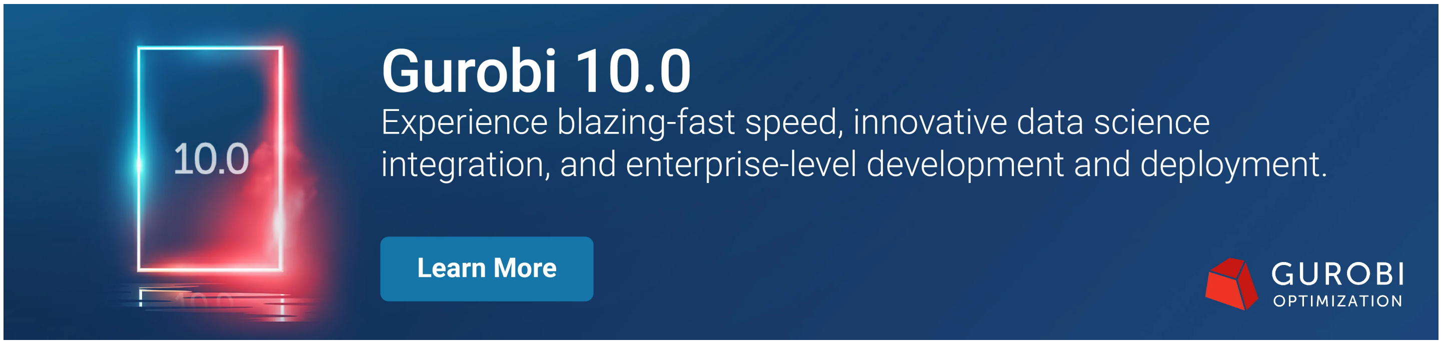


July 28, 2023 in Analyze This!

My Biggest Worry about Generative AI

A summer chat about ChatGPT

By Vijay Mehrotra

SHARE:

PRINT ARTICLE:

https://doi.org/10.1287/LYTX.2023.03.12

I am writing this column in the middle of my summer vacation away from the office. Alas, in today's hyperconnected world, it is much easier to stay connected and thus harder than ever to truly "get away." The reduced cost of cellular voice and data and the ubiquity of Wi-Fi connections are both a blessing (I am able to stay in touch with my family and colleagues better than ever) and a curse (it is increasingly difficult to turn off the noise).

Whether at home or abroad, what I can't seem to get away from is the constant buzz about ChatGPT/generative artificial intelligence (AI)/large language models. As Michael Watson points out in a recent LinkedIn post titled "AI Doom or Optimism?: The Best of Both Sides," the perspectives being put forth about where these technologies are leading us are all over the map.

On one hand, there are the pessimists. Many of the concerns center on education (a headline in *The Atlantic* just after the release of ChatGPT laments that "The College Essay Is Dead: Nobody Is Prepared for How AI Will Transform Academia") and job markets (an April 2023 article in *Wired* magazine asks "Who Will You Be After ChatGPT Takes Your Job?"). More ominously, other observers assert that generative AI has already brought us far closer to large-scale disaster than we even realize. A recent statement published on the website of the nonprofit Center on AI Safety asserts that "mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war" (it is signed by a very long list of scientists and technology industry luminaries). Last month, more than 1,000 researchers and technologists signed an open letter calling for a six-month pause on AI development because, they said, it poses profound risks to society and humanity.

Though Watson observes that it is somewhat harder to find AI optimists (or at least people willing to express their optimism publicly), his post also includes several positive testimonials as well. Most notably, earlier this month, technology pioneer and venture capitalist Marc Andreesson published a lengthy blog post titled "Why AI Will Save the World" in which he systematically offers his own responses to most of the concerns raised by the pessimists.

For the past several months, almost every day, I have been asked for my perspective on ChatGPT. Over time, I have developed a stock answer that sounds something like this:

"We have been becoming increasingly reliant on AI for a long time. Consider the web search engines and applications like Google Maps that we have become incredibly dependent on. But for the most part, both the use cases and the data used to support them are well understood. What is different about today's large language models is that they can be used to solve so many different problems – and the question that is most interesting to me is 'who will try to use these types of applications, and for what?'"

This rather anodyne response is neither controversial nor easy to argue with and has thus far allowed me to disguise how little I actually know about the technical details, challenges and risks associated with the rapidly evolving world of generative AI.

Cybercrime: The Real AI Risk?

On this vacation, however, I have come to realize that I'm far less afraid of a mass extinction event caused by sentient AI and far more concerned about these incredibly powerful multiuse tools finding their way into the hands of cybercriminals. More and more of the activity in our incredibly interconnected world is taking place online, with the COVID-19 pandemic having accelerated many existing trends. As such, even prior to the release of ChatGPT, the annual cost of cybercrime was estimated at roughly *$6 trillion* and had been projected to grow at a 15% annual rate. Thus, what really worries me about generative AI is that it could easily end up arming malicious hackers with a much more powerful set of tools, giving them the ability to disrupt our personal and professional lives in potentially catastrophic ways.

Even Andreesson readily admits that "AI will make it easier for bad people to do bad things." Not surprisingly, he suggests that we fight fire with fire:

"The same capabilities that make AI dangerous in the hands of bad guys with bad goals make it powerful in the hands of good guys with good goals … for example, if you are worried about AI generating fake people and fake videos … the answer is not to ban word processors and Photoshop – or AI – but to use technology to build a system that actually solves the problem. And so, let's mount major efforts to use AI for good, legitimate, defensive purposes."

But there is at least one major challenge in implementing this type of AI-driven defense: data. The leading generative AI systems are driven by unbelievably large data sets. ChatGPT, for example, was trained on the Common Crawl data set that includes more than 3 billion web pages. Building the types of defensive AI-driven systems that Andreesson envisions will require not only new systems and data structures but also data about past attacks and strategies that can help train those types of defensive systems.

Cybersecurity researchers have been thinking about these types of issues for some time. In their 2017 paper, "Multifirm Models of Cybersecurity Investment Competition vs. Cooperation and Network Vulnerability" (published in the *European Journal of Operational Research*), Anna Nagurney and Shivani Shukla utilize Nash bargaining theory to argue for information sharing across firms, quantifying monetary and security benefits in terms of reducing network vulnerability to cyberattacks. It is clear that the economic case for this type of information sharing and collaboration has become even stronger with the recent emergence of generative AI platforms, and there is absolutely no chance of putting that genie back in the bottle. Thus, as Nagurney and Shukla suggest, our ability to develop the types of defensive AI solutions that Andreesson imagines will depend in large part on our ability to share information effectively.

With all that said, I am going to get back to my vacation. ChatGPT – and its many opportunities and risks – will surely be waiting for me when I get back to the office.

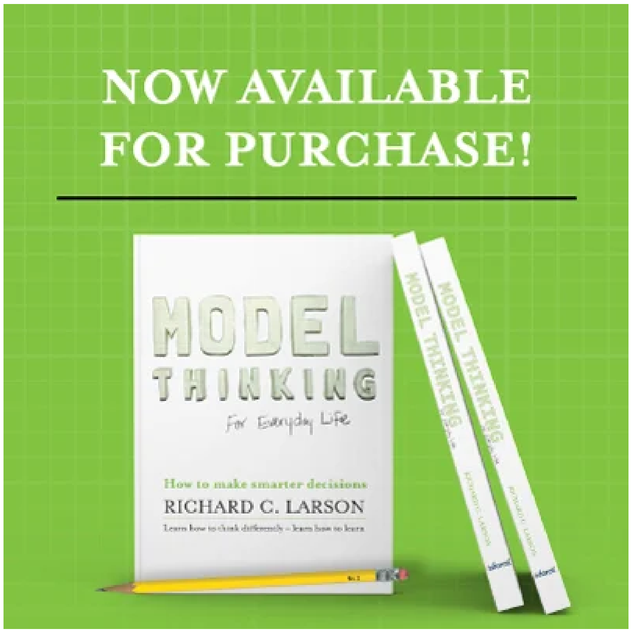
Vijay Mehrotra
(vmehrotra@usfca.edu)

Vijay Mehrotra is a professor in the Department of Business Analytics and Information Systems at the University of San Francisco’s School of Management and a longtime member of INFORMS.

SHARE:

Keywords: Analyze This!; generative AI; artificial intelligence; ChatGPT; large language models; LLMs; cybercriminals; cybercrime

10.0


NOW AVAILABLE
FOR PURCHASE!
MODEL THINKING
For Everyday Life
How to make smarter decisions
RICHARD C. LARSON
Learn how to think differently – learn how to learn

Analytics
ADVERTISE TO REACH TOP
ANALYTICS AUDIENCES!

informs
INDUSTRY OPEN FORUM

The Institute for Operations Research and the Management Sciences

5521 Research Park Drive, Suite 200
Catonsville, MD 21228 USA

phone 1 443-757-3500

phone 2 800-4INFORMS (800-446-3676)

fax 443-757-3515

email informs@informs.org