Journal of Big Data and Artificial Intelligence

WWW.JBDAI.ORG

ISSN: 2692-7977

JBDAI Vol. 2, No. 1, 2024 DOI: 10.54116/jbdai.v2i1.40

EDITORIAL

A New Era of Artificial Intelligence Begins… Where Will It Lead Us?

Jim Samuel
Bloustein, Rutgers University
jim.samuel@rutgers.edu

Abhishek Tripathi
The College of New Jersey
tripatha@tcnj.edu

Ensela Mema
Kean University
emema@kean.edu

"By far, the greatest danger of Artificial Intelligence is that people conclude too early that they understand it."
—Eliezer Yudkowsky

We have entered the age of Artificial Intelligence (AI). Everything around us is becoming artificially intelligent: from business applications to healthcare, education to finance, and governance to art, music, and entertainment. The fact that AI has gripped public attention is evident from the steep rise in public engagement with artificial intelligence applications, explosive increase in news media coverage of AI, increasing volumes of social media posts, and the mushrooming of a range of AI ecosystem initiatives. Consider the steep rise in searches for the terms "AI" and "Artificial Intelligence" from 2022 to 2023 as presented by Google Trends[1] (Figures 1 and 2, https://trends.google.com/trends/). In some senses, this is reminiscent of the big data and data-driven-everything phenomena from around two decades ago. This perspective is also supported by a review of Google Trends for search terms on big data, which shows interest in "big data" peaking around 2014. As with AI, expectations were high then, and concerns, especially around data privacy, were also significant. What many did not realize in those early years was that the big data revolution was a forerunner to what we have just begun to experience—the new era of artificial intelligence, also popularly called the fourth industrial revolution.

Undoubtedly, there have been a series of remarkable breakthroughs in AI technologies over the past few years. However, would these have been possible without big data fueling much of the early opportunities in AI (Kersting and Meyer 2018)? Big data and high-performance computing concepts have served as the foundations for the presently visible waves of AI, driven largely through the use of supervised machine learning methods, and a converse early-stage perspective has been posited with the narrative of how AI methods have been used to create value from big data (O'Leary 2013). Big data and AI have been paired not only in the realm of opportunities but also in risks and challenges

[1]Google Trends provides a relative view to gauge trends and not exact numbers—this means that the peak is treated as 100% in each figure, and figures cannot be compared except for rate of change (https://trends.google.com/trends/).

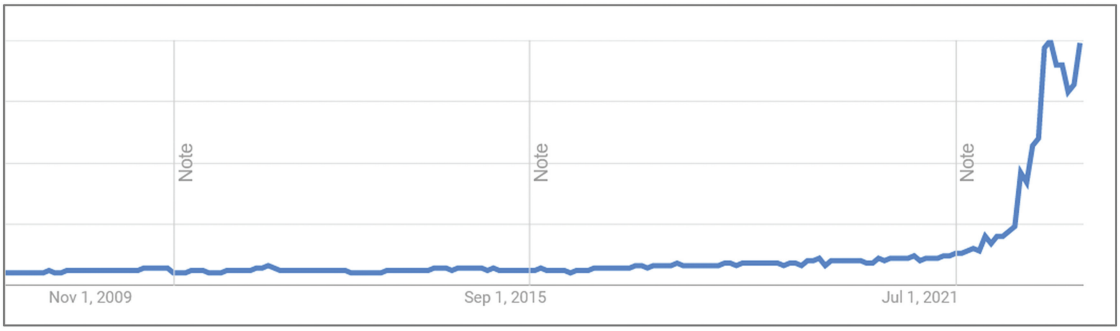

Figure 1: Google search trend for the term "AI."

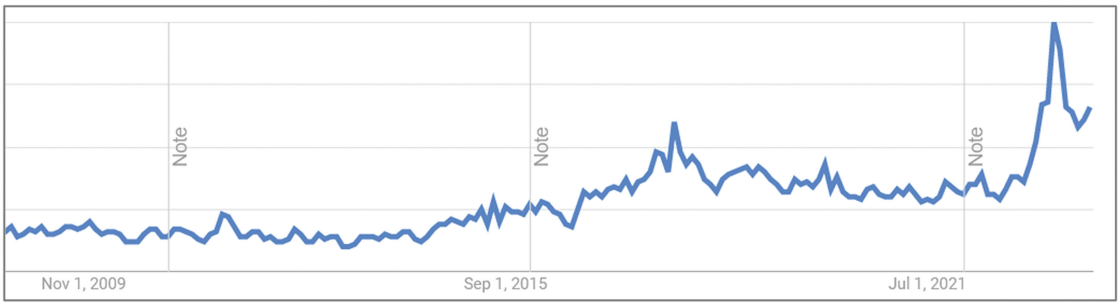

Figure 2: Google search trend for the term "Artificial Intelligence."

to human society (Helbing et al. 2019). Despite experts like Andrew Ng highlighting the challenges of big data dependence and emphasizing the emergence of "smart data"-driven AI, the fact remains that big data is here to stay as a phenomenon (Strickland 2022). This is clearly evident from recent developments in AI and the rush to develop foundation models and large language models (LLMs) based on modeling vast quantities of data, often quipped to represent "the knowledge of the world." However, "artificial intelligence" has become a broad and fuzzy term, making it "very difficult to mark its boundaries precisely and specify what exactly it encompasses" (Devedzic 2022).

What is artificial intelligence? A simple yet encompassing definition expands on Samuel (2021) and Samuel et al. (2022): "Artificial Intelligence is a set of technologies that mimic the functions and expressions of human intelligence, specifically cognition, logic, learning, adaptivity and creativity." Auxiliary topics such as AI risks and AI ethics need to be defined separately as they arise from the potential impacts of AI. There is a need for continued research on various dimensions of AI. As the domain matures, it becomes imperative to develop robust AI frameworks from conceptual, ontological, and "epistemological, philosophical, ethical, technological, and social perspectives" (Devedzic 2022). The era of AI has begun with an explosive expansion that is rapidly influencing nearly every facet of human life and society—it is crucial to recognize that the current wave of AI applications is driven by developments with little philosophical or moral foundations for human interaction with advanced technologies that are simultaneously supersmart and superstupid (absence of consistent commonsense in AI). This is reflected in a number of ethics and transparency concerns surrounding the development of foundation models and LLM applications (Samuel 2023). One of the greatest present needs is the development of a robust philosophy of AI that includes interaction with human intelligence.

AI applications in major categories such as natural language processing (NLP), computer vision (CV), intelligent robotics, intelligent and connected sensors and smart quantitative models, have all used machine learning [mostly supervised, unsupervised and reinforcement learning (RL)] and some form of big data. Even though RL can operate with smaller simulated data, it has been successfully applied to big data. Consider the recent release of state-of-the-art AI foundation models such as PaLM (March 2023, with 540 billion parameters), LLaMA-2 [now freely available for downloads in multiple sizes (LLaMA-2 2023; LLaMA 2-Meta 2023)], and GPT-4 (and GPT-5 in the works), and open-source models such as BLOOM. Generative AI has been researched for over two decades, and extant research has explored generative concepts and applied it to areas such as social media (Sutskever et al. 2011; Reed et al. 2016; Garvey et al. 2021). However, the late 2022 launch of the GPT-3+-based ChatGPT application brought awareness about the power and the usefulness of generative AI to the forefront. Google (Alphabet), as an

example, has continued to release new models and applications—Chirp is a cluster of "Universal Speech Models" built on 12 million hours of speech data for automatic speech recognition (ASR); Imagen, Muse and Parti are different types of text-to-image models, while Codey drafts code for humans (Google AI 2023). While we are still absorbing these, in addition to PaLM-2, Alphabet CEO Sundar Pichai announced Google's next superfoundation-model "Gemini," which is reportedly still in training (Pichai 2023). These and similar developments herald the new era of artificial intelligence—it will be wise for us to assume that our understanding is limited and we need substantial research and thought leadership on big data and AI in the years ahead to harness AI effectively for the benefit of human society.

Welcome to the Journal of Big Data and Artificial Intelligence! Keeping in line with this overarching trend towards AI-driven value creation, and the critical need to further catalyze research in artificial intelligence across multiple dimensions and domains, the Executive Editorial Board of this journal has decided to affirm this undeniably increasing importance of AI research in the context of big data: Starting with this volume, the *Journal of Big Data Theory and Practice* has been formally renamed as the *Journal of Big Data and Artificial Intelligence* (***JBDAI***). This move was deliberated over for a significant period of time and went through multiple evaluation cycles and ratification with the parent body, New Jersey Big Data Alliance (NJBDA). The *Journal of Big Data and Artificial Intelligence* now better represents the research it publishes on big data, machine learning, NLP, domain analytics, image processing, LLMs, and other AI research. We also anticipate a significant enhancement of the brand value of the journal and an increase in the attractiveness of the journal to a broader range of researchers. Please join us in this venture to create scholarly thought leadership in big data and AI by contributing at various levels to *JBDAI*.

This volume presents five intellectually stimulating articles, a memorial to one of our key scholars and founding member, and this present editorial—these articles traverse a broad range of topics and methods and we hope along with the authors that these will be useful: The "BERT based Blended approach for Fake News Detection" paper provides an interesting approach to identifying fake news on social media using BERT-LSTM and BERT-CNN. The "Investment under Uncertainty: The Role of Inventory Dynamics" paper demonstrates a novel approach to study the influence of inventory on the value of the firm and on investment decisions. Authors of the "Crime Frequency During Covid-19 and Black Lives Matter Protests" use the Holt-Winters and SARIMA models to explore changes to crime under the pandemic and social unrest conditions. Applying image classification methods, the "Machine Learning Study: Identification of Skin Diseases for Various Skin Types Using Image Classification" paper shows an approach to improving classification accuracy. Finally, the "Are Emotions Conveyed Across Machine Translations? Establishing an Analytical Process for the Effectiveness of Multilingual Sentiment Analysis with Italian Text" uses natural language processing methods within a multilingual context to generate a repeatable process for analyzing text.

In conclusion, it is necessary to highlight the irreversible connection between big data and AI—each of the above mentioned (and more) foundation models and AI applications are built on vast quantities of data, and the multimodal foundation models are each built on multiple types of big data such as text, audio, and images. Undoubtedly, we will witness the development of AIs trained on relatively smaller quantities of "smart data," and other variations. However, we are yet to observe any significant developments signaling the replacement of big data without compromising efficacy. It is more likely that smart data, small data, and all such variations to locally defined data-size optimality will eventually add up leading to a continued net increase in big data volumes and complexities. Big data and AI are, from all that we can presently observe, irreversibly linked together, with AI methods being used to generate insights and value from big data, and big data serving as foundational building blocks for a vast array of AI applications. Furthermore, AI research is rapidly expanding in the social sciences, cultural studies, socioeconomics and education, among other domains (Samuel et al. 2023). We at *JBDAI* hope to encourage and foster much high-quality research, rigor, and innovative thought leadership on big data and artificial intelligence in the years ahead, supporting human well-being, the sustainability of our natural resources, and balanced societal progress—please contribute to *JBDAI* and be a part of this exciting intellectual adventure!

References

Devedzic, V. 2022. "Identity of AI." *Discover Artificial Intelligence* **2**, no. 1: 23. doi: 10.1007/s44163-022-00038-0

Garvey, M. D., J. Samuel, and A. Pelaez. 2021. "Would You Please like my Tweet?! an Artificially Intelligent, Generative Probabilistic, and Econometric Based System Design for Popularity-Driven Tweet Content Generation." *Decision Support Systems* **144**: 113497. doi: 10.1016/j.dss.2021.113497

Google AI. 2023. "Foundation Models." Accessed November 27, 2023. https://ai.google/discover/foundation-models/

Helbing, D., B. S. Frey, G. Gigerenzer, E. Hafen, M. Hagner, Y. Hofstetter, J. van den Hoven, R. V. Zicari, and A. Zwitter. 2019. "Will Democracy Survive Big Data and Artificial Intelligence?" In *Towards Digital Enlightenment: Essays on the Dark and Light Sides of the Digital Revolution*, edited by Dirk Helbing, 73–98. Springer: Cham, Switzerland. doi: 10.1007/978-3-319-90869-4_7

Kersting, K., and U. Meyer. 2018. "From Big Data to Big Artificial Intelligence? Algorithmic Challenges and Opportunities of Big Data." *KI - Künstliche Intelligenz* **32**, no. 1: 3–8. doi: 10.1007/s13218-017-0523-7

LLaMA-2. 2023. "Model Download." Accessed November 27, 2023. https://huggingface.co/meta-llama

LLaMA 2-Meta. 2023. "Llama 2: Open Foundation and Fine-Tuned Chat Models." Accessed November 27, 2023. https://ai.meta.com/research/publications/llama-2-open-foundation-and-fine-tuned-chat-models/

O'Leary, D. E. 2013. "Artificial Intelligence and Big Data." *IEEE Intelligent Systems* **28**, no. 2: 96–99. doi: 10.1109/MIS.2013.39

Pichai, S. 2023. "Google I/O 2023: Making AI More Helpful for Everyone." Accessed November 27, 2023. https://blog.google/technology/ai/google-io-2023-keynote-sundar-pichai/#ai-products

Reed, S., Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee. 2016. "Generative Adversarial Text to Image Synthesis." In *International Conference on Machine Learning*, June 19–24, New York, USA.

Samuel, J. 2021. "A Call for Proactive Policies for Informatics and Artificial Intelligence Technologies." Scholars Strategy Network. Accessed November 27, 2023. https://scholars.org/contribution/call-proactive-policies-informatics-and

Samuel, J. 2023. "The Critical Need for Transparency and Regulation amidst the Rise of Powerful Artificial Intelligence Models." Scholars Strategy Network (SSN) Key Findings. Accessed November 27, 2023. https://scholars.org/contribution/critical-need-transparency-and-regulation

Samuel, J., R. Kashyap, Y. Samuel, and A. Pelaez. 2022. "Adaptive Cognitive Ft: Artificial Intelligence Augmented Management of Information Facets and Representations." *International Journal of Information Management* **65**: 102505. doi: 10.1016/j.ijinfomgt.2022.102505

Samuel, Y., Brennan-Tonetta, M., Samuel, J., Kashyap, R., Kumar, V., Madabhushi, S. K. K., Chidipothu, N., Anand, I., and Jain, P. 2023. "Cultivation of Human Centered Artificial Intelligence: Culturally Adaptive Thinking in Education for AI (CATE-AI)." *Frontiers in Artificial Intelligence* **6**. https://www.frontiersin.org/articles/10.3389/frai.2023.1198180

Strickland, E. 2022. "Andrew Ng, AI Minimalist: The Machine-Learning Pioneer Says Small is the New Big." *IEEE Spectrum* **59** (4), 22–50.

Sutskever, I., J. Martens, and G. E. Hinton. 2011. "Generating Text with Recurrent Neural Networks." In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, Washington, USA, June 28–July 2.