PERSPECTIVES
VIEWPOINTS FROM THE ACADIAN TEAM

Generative AI in Systematic Investing: The Sizzle and the Steak

April 2024

- Generative AI gained widespread attention following the debut of ChatGPT by OpenAI in 2022. While the technology may seem new, its origins and applications trace back to the mid-2000s.

- Even though it's their capacity to produce new realistic content that garners both wonder and concern, the significance and potential impact of Generative AI models extend far beyond their purely generative use cases.

- The source of generative AI's power is also its limitation. The vast amounts of data such models require for training may not exist for many investment problems. Knowing where, when, and how to successfully apply Generative AI to systematic investing will be a source of competitive edge.

Generative AI: Hardly Out of Nowhere

While the barcode was invented in the late 1940s, most people were unaware of the technology's existence prior to its widespread adoption in supermarkets decades later. Checkout scanners and their ubiquitous beeps rapidly proceeded to become so engrained in daily life that President George H. W. Bush's apparent (but misreported) unfamiliarity with the technology at a campaign stop in Florida fueled a narrative that he was disconnected from ordinary voters, likely contributing to his defeat by the up-and-coming challenger, Bill Clinton.[1]

The idea of interconnected computer networks had been brewing in academic and research circles for decades, with notable developments like ARPANET in the late 1960s and Tim Berners-Lee's proposal for the World Wide Web in 1989. However, it was the release of Netscape Navigator in December 1994 that made the internet accessible to the public, leading to a rapid expansion of online services and the birth of the dot-com boom.

**Figure 1: Worldwide Search Activity for "Generative AI"**
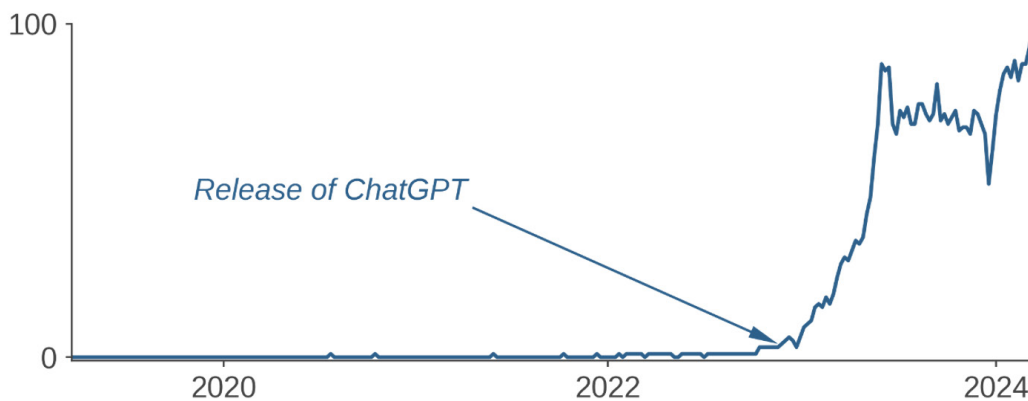
Source: Acadian based on data from Google Trends. Reflects worldwide search activity for the term "Generative AI." For illustrative purposes only.

---

[1] "AP Was There: Bush's Bum Rap on 'Amazing' Barcode Scanner," APNews.com, December 4, 2018.

For institutional investor use only. Not to be reproduced or disseminated.

1

SIDEBAR: Generative AI in Broader Context

ARTIFICIAL INTELLIGENCE (AI)

Mimicking the intelligence or behavioral pattern of humans or any other living entity.
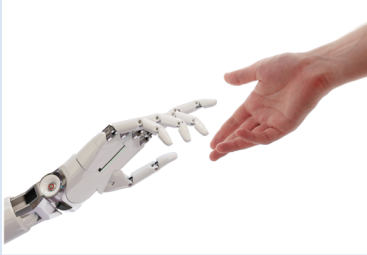
MACHINE LEARNING (ML)

A technique by which a computer can "learn" from data, without using a complex set of different rules. This approach is mainly based on training a model from datasets.

DEEP LEARNING

A technique to perform machine learning inspired by the human brain's own network of neurons.
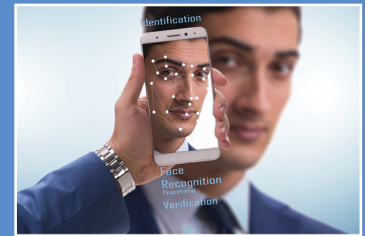
The field of artificial intelligence got its spark in 1950 with the publication of Alan Turing's seminal paper, "Computing Machinery and Intelligence."* It was born as an academic discipline at the Dartmouth Conference, organized by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon. Nascent AI efforts in the 1950s followed and focused on simple rules-based systems. Over the subsequent decades, researchers attempted to replicate human expertise in various specific domains with limited success. The field experienced periods of optimism, marked by significant progress and excitement. These were punctuated by "AI winters," characterized by funding cuts and decreased interest due to unmet expectations and limited progress. AI applications continued to grow over time but remained basic due to limitations in training data and computing power.

Over the past 20 years, exponential growth in volumes of data and computing power permitted significant increases in AI sophistication and capabilities. Methodological breakthroughs were just as important to progress. One of these was Ian Goodfellow's eureka moment during a casual gathering with friends at a Montreal bar in 2014. What if you pitted two neural networks against each other, he wondered, one generating synthetic data and the other discriminating whether that output is real or fake? The adversarial dynamic between them would eventually converge when the generator creates output so realistic that the discriminator can no longer tell it apart from the real thing.

This groundbreaking idea gave rise to Generative Adversarial Networks (GANs), which have since revolutionized many areas of machine learning, opening up new frontiers in data generation, image synthesis, and beyond. GANs, Transformers, and many other innovations have thrust the field forward, enabling machines to perform increasingly complex tasks, matching and at times exceeding human capabilities.

---

* Turing, Alan M., "Computing Machinery and Intelligence." Mind 59, issue 236, (1950): 433-460.

For institutional investor use only. Not to be reproduced or disseminated.

2

These are not isolated phenomena. A technology often emerges and makes strides for years largely out of public view before capturing mass attention with a prominent milestone or product release, leading to the misconception that it was completely new. Generative AI, a branch of artificial intelligence so named for models that are trained to "generate" new data, has followed a similar path to mainstream consciousness.[2]

While generative AI surged into the zeitgeist seemingly out of nowhere in late 2022 with OpenAI's release of ChatGPT (Figure 1), this subfield of artificial intelligence has been around since at least the mid-2000s. If you have ever used Google Translate, Autocomplete, Siri, Alexa or come across a deepfake image or video, then you have seen Generative AI in action.[3] While excitement about generative AI is certainly warranted, we find that current speculations about its applicability and promise are in some cases misplaced and in others premature – both broadly speaking as well as in the context of systematic investing.

Legitimate Excitement

The output of generative models has been increasing in sophistication and realism at a truly remarkable pace. While it is their generative applications that have received much of the attention, we believe their potential benefits are dwarfed by those of applications that don't require creation of new content but where Generative AI can nonetheless excel relative to simpler models – provided certain conditions are met (more on this below).

In active systematic investing, strictly generative use cases are fairly limited – missing data imputation, synthetic data generation, and risk modelling are some of the more frequently mentioned examples. Generative applications of Natural Language Processing (NLP) based on large language models could increase operating efficiency in other parts of the business. The non-generative applications, on the other hand, are potentially far more numerous and could theoretically augment or even replace many of the existing methods employed in modeling alpha, risk, and transaction costs. It is this side of generative AI that we find far more exciting. In order to see why, we need to better understand where generative AI draws its strengths.

Generative models excel at capturing high-dimensional data distributions. To achieve realistic synthesis, these models need to learn and reproduce the intricate structures and relationships present in the data, whether images, text, or other complex information. Put simply, in order to generate something, one needs to understand it on a much deeper level than only to recognize or classify it.[4] All of us can tell a cat from a dog, but how many can draw a realistic rendering of either? A generative model able to synthetically construct such images will, after additional tuning, perform better at classifying them than a simpler model trained to perform classification only. Similarly, categorizing excerpts of text based on their sentiment is a much simpler task than that of writing a happy or a melancholy poem based on a simple prompt. Understanding sentiment involves grasping the context in which certain words or phrases appear. Generative models excel in capturing subtle contextual nuances to achieve their semantic understanding.[5] Thus, a generative model should achieve superior results in extracting sentiment or in performing other tasks that require understanding the contents of a given body of text.

To summarize, the key advantage of generative models is that because they are constructed to perform the more ambitious task of creating new content, they can perform better at non-generative tasks than simpler models designed only for those less sophisticated purposes. This significantly increases generative AI's scope of applicability.

Theory Versus Practice

Generative AI's strengths come with limitations and challenges. But these issues are too dry or complex to garner much attention in wide-eyed popular media coverage.

The models are huge, some containing billions of parameters (or more). As a result, training generative models, i.e., estimating those parameters, requires enormous computational resources and massive amounts of data. While the main obstacle in bringing to bear the requisite computing power has become the price tag, data will, in many contexts, pose a barrier that is insurmountable: the volumes of data required to apply generative AI to some problems may simply not yet exist. Active investing is

---

[2] At their core, all AI techniques represent algorithms leveraging patterns identified in real-world data. For many years, they have been used in an increasingly broad array of disciplines, including systematic investing. Over time, data scientists have developed progressively more sophisticated AI methods, called deep learning, i.e., complex neural network-based systems that emulate how the human brain performs pattern recognition. Generative AI is an application of deep learning with models designed and trained to produce entirely new content, including images, text, and other complex data.

[3] Restricted Boltzmann Machines (RBMs) were used to model user preferences in recommendation systems as early as 2006. Google Translate was launched the same year, followed by Autocomplete four years later. Siri was integrated into iPhones in 2011. Although variational autoencoders could generate and reconstruct images even earlier, Ian Goodfellow's invention of Generative Adversarial Networks (GANs) was a major methodological breakthrough, greatly increasing their realism and leading to numerous applications – deepfakes being the more notorious among them. Another significant achievement occurred when Google introduced the Transformer architecture in 2017, which has since become a cornerstone in language modeling and serves as the foundation for ChatGPT.

[4] Traditional AI models are termed discriminative because they produce distinct classifications or predictions based on identified patterns. The algorithms recognize patterns either 1) after being trained to do so with example data that contains known outcomes (supervised learning) or 2) autonomously without explicit guidance or predefined solutions (unsupervised learning).

[5] These benefits aren't simply due to the size of deep neural networks and the massive datasets they are trained on. While both generative and non-generative models can benefit from large training datasets and deep neural networks, there are specific advantages that the generative component offers, contributing to a potentially deeper understanding of language intricacies. Generative models, especially those based on transformer architectures like GPT, are designed to better capture contextual information. They can learn implicit features that are not explicitly labeled in the training data. This ability allows them to capture subtle patterns and relationships associated with sentiment that might be challenging for non-generative models and to achieve a certain level of abstraction and comprehension. This creative aspect can contribute to a more holistic understanding of language and sentiment. Please contact us for further discussion of this point.

replete with cases where available data sets are nowhere near the scale needed to train data-hungry generative models. Data availability, in fact, helps to explain why NLP has been one area of focus—there are massive corpora of written and spoken text with which researchers can train sophisticated models before applying them to domain-specific projects. Managers must decide whether to rely on an off-the-shelf "foundation model," i.e., one that is pre-trained based on publicly available information, or to build a customized model trained from scratch on their own data better reflective of that use case. The former may provide acceptable performance for some applications, such as in quantifying characteristics of text that does not require a lot of domain knowledge to parse or interpret. But where specialized knowledge is a must, generic models will struggle without additional tuning.[6] For example, as of this writing, ChatGPT does not recommend itself as a tool to quantify sentiment of earnings call transcripts and instead "humbly" suggests better-suited simpler methods.[7]

Peculiarities of the active investing context accentuate the importance of design and methodological choices in the application of any machine-learning techniques, including generative AI. The signal-to-noise ratio in forecasting asset returns (and/or various related KPIs) tends to be significantly lower than in many other predictive contexts, the features and rules of the market environment are in constant flux, and players rapidly adapt their behavior in response to other participants. Such characteristics put a premium on judicious engineering, including methods of finding the "right fit" (to avoid both over- and under-fitting), design of input features, optimal use of the available data, etc. Successful application of generative AI in systematic investing, therefore, requires significant investment in specialized expertise.

Conclusion

Generative AI represents a powerful addition to the investor's toolkit. It is not an utterly new development nor does its main promise lie in its generative functionality. It offers systematic managers new capabilities as well as the potential to improve many of their existing quantitative models. Yet the massive amounts of data required to fuel generative AI's power also limit its application. In systematic investing, knowing where, when, and how to successfully apply generative AI will demand a combination of model engineering expertise and financial domain knowledge that will represent another source of competitive edge.

---

[6] E.g., language models can suffer from hallucination and knowledge limitation, including outdated or incorrect facts, and lack of specialized knowledge. This is why knowledge augmentation approaches such as Retrieval Augmentation Generation (RAG)/Langchain have gained in popularity.

[7] E.g., suggesting several variants of BERT (Bidirectional Encoder Representation from Transformers) – an NLP model which uses bidirectional context to understand text.

BIOGRAPHIES

Vladimir Zdorovtsov, Ph.D.
SENIOR VICE PRESIDENT, DIRECTOR, GLOBAL EQUITY RESEARCH

Vlad joined Acadian in 2019 and is Director of Global Equity Research. Previously, he was a managing director at Goldman Sachs Asset Management (GSAM), co-leading a quantitative equity research team within the firm's fundamental equity and direct alternatives units. In this role, Vlad's focus was on alpha capture and the integration of quantitative alpha, risk, and portfolio construction techniques. Prior to that, he was the head of research at Rosetta Analytics, Inc., a machine learning hedge fund. Vlad received a Ph.D. in finance from the University of South Carolina and holds an M.B.A. from the University of Arkansas as well as a B.A. in international economics from Sumy State University, Ukraine.

Anshuman Ramachandran
ASSISTANT VICE PRESIDENT, ASSOCIATE INVESTMENT STRATEGIST

Anshuman is an Associate Investment Strategist in Acadian's Client Advisory team. He works on research projects that are driven by interest from Acadian's investor base, conducting original empirical analysis and communicating findings. He has also driven development of the team's analytical resources. Prior to joining Acadian full time in 2021, he was an intern with Acadian's Equity Alternative Strategies team. He holds an M.S. in financial engineering from Columbia University and his undergraduate degree in mechanical engineering from IIT Madras.

Seth Weingram, Ph.D.
SENIOR VICE PRESIDENT, DIRECTOR, CLIENT ADVISORY

Seth is Acadian's Director of Client Advisory, a team that produces original research on topical issues in systematic investing and regularly meets with key clients, consultants, and prospects. Seth also chairs Acadian's Editorial Board, driving the firm's thought leadership. Prior to joining Acadian in 2014, Seth held senior roles in equity derivatives trading, research, and strategy at UBS, Barclays Global Investors, and Deutsche Bank. Seth holds a Ph.D. in economics from Stanford University and a B.A. in economics from the University of Chicago.

General Legal Disclaimer

These materials provided herein may contain material, non-public information within the meaning of the United States Federal Securities Laws with respect to Acadian Asset Management LLC, BrightSphere Investment Group Inc. and/or their respective subsidiaries and affiliated entities. The recipient of these materials agrees that it will not use any confidential information that may be contained herein to execute or recommend transactions in securities. The recipient further acknowledges that it is aware that United States Federal and State securities laws prohibit any person or entity who has material, non-public information about a publicly-traded company from purchasing or selling securities of such company, or from communicating such information to any other person or entity under circumstances in which it is reasonably foreseeable that such person or entity is likely to sell or purchase such securities.

Acadian provides this material as a general overview of the firm, our processes and our investment capabilities. It has been provided for informational purposes only. It does not constitute or form part of any offer to issue or sell, or any solicitation of any offer to subscribe or to purchase, shares, units or other interests in investments that may be referred to herein and must not be construed as investment or financial product advice. Acadian has not considered any reader's financial situation, objective or needs in providing the relevant information.

The value of investments may fall as well as rise and you may not get back your original investment. Past performance is not necessarily a guide to future performance or returns. Acadian has taken all reasonable care to ensure that the information contained in this material is accurate at the time of its distribution, no representation or warranty, express or implied, is made as to the accuracy, reliability or completeness of such information.

This material contains privileged and confidential information and is intended only for the recipient/s. Any distribution, reproduction or other use of this presentation by recipients is strictly prohibited. If you are not the intended recipient and this presentation has been sent or passed on to you in error, please contact us immediately. Confidentiality and privilege are not lost by this presentation having been sent or passed on to you in error.

Acadian's quantitative investment process is supported by extensive proprietary computer code. Acadian's researchers, software developers, and IT teams follow a structured design, development, testing, change control, and review processes during the development of its systems and the implementation within our investment process. These controls

annual independent review by our SOC1 auditor. However, despite these extensive controls it is possible that errors may occur in coding and within the investment process, as is the case with any complex software or data-driven model, and no guarantee or warranty can be provided that any quantitative investment model is completely free of errors. Any such errors could have a negative impact on investment results. We have in place control systems and processes which are intended to identify in a timely manner any such errors which would have a material impact on the investment process.

Acadian Asset Management LLC has wholly owned affiliates located in London, Singapore, and Sydney. Pursuant to the terms of service level agreements with each affiliate, employees of Acadian Asset Management LLC may provide certain services on behalf of each affiliate and employees of each affiliate may provide certain administrative services, including marketing and client service, on behalf of Acadian Asset Management LLC.

Acadian Asset Management LLC is registered as an investment adviser with the U.S. Securities and Exchange Commission. Registration of an investment adviser does not imply any level of skill or training.

Acadian Asset Management (Singapore) Pte Ltd, (Registration Number: 199902125D) is licensed by the Monetary Authority of Singapore. It is also registered as an investment adviser with the U.S. Securities and Exchange Commission.

Acadian Asset Management (Australia) Limited (ABN 41 114 200 127) is the holder of Australian financial services license number 291872 ("AFSL"). It is also registered as an investment adviser with the U.S. Securities and Exchange Commission. Under the terms of its AFSL, Acadian Asset Management (Australia) Limited is limited to providing the financial services under its license to wholesale clients only. This marketing material is not to be provided to retail clients.

Acadian Asset Management (UK) Limited is authorized and regulated by the Financial Conduct Authority ('the FCA') and is a limited liability company incorporated in England and Wales with company number 05644066. Acadian Asset Management (UK) Limited will only make this material available to Professional Clients and Eligible Counterparties as defined by the FCA under the Markets in Financial Instruments Directive, or to Qualified Investors in Switzerland as defined in the Collective Investment Schemes Act, as applicable.

ACADIAN

GLOBAL AFFILIATES

Boston   London   Singapore   Sydney

ACADIAN-ASSET.COM