

THE CAPCO INSTITUTE
JOURNAL
OF FINANCIAL TRANSFORMATION

ORGANIZATIONAL

Generative artificial intelligence
assessed for asset management

UDO MILKAU

**ARTIFICIAL
INTELLIGENCE**

#58 NOVEMBER 2023

a **wipro** company

THE CAPCO INSTITUTE

JOURNAL OF FINANCIAL TRANSFORMATION

RECIPIENT OF THE APEX AWARD FOR PUBLICATION EXCELLENCE

Editor

Shahin Shojai, Global Head, Capco Institute

Advisory Board

Michael Ethelston, Partner, Capco

Farzine Fazel, Partner, Capco

Anne-Marie Rowland, Partner, Capco

Editorial Board

Franklin Allen, Professor of Finance and Economics and Executive Director of the Brevan Howard Centre, Imperial College London and Professor Emeritus of Finance and Economics, the Wharton School, University of Pennsylvania

Philippe d'Arvisenet, Advisor and former Group Chief Economist, BNP Paribas

Rudi Bogni, former Chief Executive Officer, UBS Private Banking

Dan Breznitz, Munk Chair of Innovation Studies, University of Toronto

Elena Carletti, Professor of Finance and Dean for Research, Bocconi University, Non-Executive Director, Unicredit Spa

Lara Cathcart, Associate Professor of Finance, Imperial College Business School

Jean Dermine, Professor of Banking and Finance, INSEAD

Douglas W. Diamond, Merton H. Miller Distinguished Service Professor of Finance, University of Chicago

Elroy Dimson, Emeritus Professor of Finance, London Business School

Nicholas Economides, Professor of Economics, New York University

Michael Enthoven, Chairman, NL Financial Investments

José Luis Escrivá, President, The Independent Authority for Fiscal Responsibility (AIReF), Spain

George Feiger, Pro-Vice-Chancellor and Executive Dean, Aston Business School

Gregorio de Felice, Head of Research and Chief Economist, Intesa Sanpaolo

Maribel Fernandez, Professor of Computer Science, King's College London

Allen Ferrell, Greenfield Professor of Securities Law, Harvard Law School

Peter Gomber, Full Professor, Chair of e-Finance, Goethe University Frankfurt

Wilfried Hauck, Managing Director, Statera Financial Management GmbH

Pierre Hillion, The de Picciotto Professor of Alternative Investments, INSEAD

Andrei A. Kirilenko, Professor of Finance, Cambridge Judge Business School, University of Cambridge

Katja Langenbucher, Professor of Banking and Corporate Law, House of Finance, Goethe University Frankfurt

Mitchel Lenson, Former Group Chief Information Officer, Deutsche Bank

David T. Llewellyn, Professor Emeritus of Money and Banking, Loughborough University

Eva Lomnicka, Professor of Law, Dickson Poon School of Law, King's College London

Donald A. Marchand, Professor Emeritus of Strategy and Information Management, IMD

Colin Mayer, Peter Moores Professor of Management Studies, Oxford University

Francesca Medda, Professor of Applied Economics and Finance, and Director of UCL Institute of Finance & Technology, University College London

Pierpaolo Montana, Group Chief Risk Officer, Mediobanca

John Taysom, Visiting Professor of Computer Science, UCL

D. Sykes Wilford, W. Frank Hipp Distinguished Chair in Business, The Citadel

CONTENTS

TECHNOLOGICAL

08 Overview of artificial intelligence deployment options

Ali Hirsa, Professor of Professional Practice, Department of Industrial Engineering and Operations Research, Columbia University, and Chief Scientific Officer, ASK2.AI

Satyan Malhotra, Chief Executive Officer, ASK2.AI

24 Applied generative AI governance: A viable model through control automation

Gerhardt Scriven, Managing Principal

Marcel Braga, Principal Consultant

Diogo Santos, Principal Consultant

Diego Sarai, Managing Principal

34 AI and banks. In conversation with an AI intern

Jesús Lozano Belio, Senior Manager, Digital Regulation, Regulation and Internal Control, BBVA

44 Performance of using machine learning approaches for credit rating prediction: Random forest and boosting algorithms

W. Paul Chiou, Associate Teaching Professor of Finance, Northeastern University

Yuchen Dong, Senior Engineer, MathWorks

Sofia X. Ma, Senior Engineer, MathWorks

54 A smart token model for native digital assets

Ian Hunt, Buy-Side Industry Consultant and Adviser

OPERATIONAL

72 Networked business design in the context of innovative technologies: Digital transformation in financial business ecosystems

Dennis Vetterling, Doctoral candidate, Institute of Information Management, University of St. Gallen

Ulrike Baumöl, Executive Director of Executive Master of Business Administration in Business Engineering, and Senior Lecturer on Business Transformation, University of St. Gallen

82 Developers 3.0: Integration of generative AI in software development

Fayssal Merimi, Managing Principal, Capco

Julien Kokocinski, Partner, Capco

90 Digital transformation and artificial intelligence in organizations

Niran Subramaniam, Associate Professor in Financial Management & Systems, Henley Business School

98 Is accounting keeping pace with digitalization?

Alnoor Bhimani, Professor of Management Accounting and Director of the South Asia Centre, London School of Economics

104 Bank and fintech for transformation of financial services: What to keep and what is changing in the industry

Anna Omarini, Tenured Researcher, Department of Finance, Bocconi University

ORGANIZATIONAL

116 The truth behind artificial intelligence: Illustrated by designing an investment advice solution

Claude Diderich, Managing Director, innovate.d

126 Duty calls – but is industry picking up?

Jessica Taylor, Consultant, Capco

Ivo Vlaev, Professor of Behavioral Science, Warwick Business School

Antony Elliott OBE, Founder, The Fairbanking Foundation

138 Generative artificial intelligence assessed for asset management

Udo Milkau, Digital Counsellor

150 How can banks empower their customers to flag potential vulnerabilities?

Przemek de Skuba, Senior Consultant, Capco

Bianca Gabellini, Consultant, Capco

Jessica Taylor, Consultant, Capco

160 Assessing AI and data protection expertise in academia and the financial services sector: Insights and recommendations for AI skills development

Maria Moloney, Senior Researcher and Consultant, PrivacyEngine, Adjunct Research Fellow, School of Computer Science, University College Dublin

Ekaterina Svetlova, Associate Professor, University of Twente

Cal Muckley, Professor of Operational Risk in the Banking and Finance Area, UCD College of Business, and Fellow, UCD Geary Institute

Eleftheria G. Paschalidou, Ph.D. Candidate, School of Economics, Aristotle University of Thessaloniki

Ioana Coita, Consultant Researcher, Faculty of Economics, University of Oradea

Valerio Poti, Professor of Finance, Business School, University College Dublin, and Director, UCD Smurfit Centre for Doctoral Research



DEAR READER,

As the financial services industry continues to embrace transformation, advanced artificial intelligence models are already being utilized to drive superior customer experience, provide high-speed data analysis that generates meaningful insights, and to improve efficiency and cost-effectiveness.

Generative AI has made a significant early impact on the financial sector, and there is much more to come. The highly regulated nature of our industry, and the importance of data management mean that the huge potential of AI must be harnessed effectively – and safely. Solutions will need to address existing pain points – from knowledge management to software development and regulatory compliance – while also ensuring institutions can experiment and learn from GenAI.

This edition of the Capco Journal of Financial Transformation examines practical applications of AI across our industry, including banking and fintechs, asset management, investment advice, credit rating, software development and financial ecosystems. Contributions to this edition come from engineers, researchers, scientists, and business executives working at the leading edge of AI, as well as the subject matter experts here at Capco, who are developing innovative AI-powered solutions for our clients.

To realize the full benefits of artificial intelligence, business leaders need to have a robust AI governance model in place, that meets the needs of their organizations while mitigating the risks of new technology to trust, accuracy, fairness, inclusivity, and intellectual property. A new generation of software developers who place AI at the heart of their approach is also emerging. Both GenAI governance and these ‘Developers 3.0’ are examined in this edition.

This year Capco is celebrating its 25th anniversary, and our mission remains as clear today as a quarter century ago: to simplify complexity for our clients, leveraging disruptive thinking to deliver lasting change for our clients and their customers. By showcasing the very best industry expertise, independent thinking and strategic insight, our Journal is our commitment to bold transformation and looking beyond the status quo. I hope you find the latest edition to be timely and informative.

Thank you to all our contributors and readers.

A handwritten signature in black ink, appearing to read 'Lance Levy', with a stylized, flowing script.

Lance Levy, **Capco CEO**

GENERATIVE ARTIFICIAL INTELLIGENCE ASSESSED FOR ASSET MANAGEMENT¹

UDO MILKAU | Digital Counsellor

ABSTRACT

Since the launch of the generative artificial intelligence tool ChatGPT end of 2022, there has been an incredible public awareness. Doomers predicted an end to humanity, while more reasonable assessments discussed the impact on traditional industries and on the workforce. In a nutshell, generative artificial intelligence is nothing more than statistical estimation and continuation of an input sequence based on a text corpus of the past. To evaluate the actual impact of generative artificial intelligence (AI) and large language models (LLM), this paper uses the case of asset management as a benchmark. These statistical estimators can produce a “next best token” based on an “internet average”, i.e., tremendous text corpora gathered from internet sources, but which cannot understand, predict anything new, or create something innovative. Consequently, generative AI/LMMs can augment staff to perform “on average”, or help internet users obtain an “average answer” to their questions about financial management. While this can (and probably will) change the future landscape of financial advice and the way consumers access information, generative AIs/LMMs are far from any type of “superintelligence”. The potential of misuse by human actors, however, remains the biggest danger and has to be monitored closely.

1. INTRODUCTION: SHOULD WE FEAR GENERATIVE ARTIFICIAL INTELLIGENCE?

Generative artificial intelligence (AI) is a topic of much debate. Between March and May 2023, three open letters [Clarkson Law Firm (2023)] have been published about the existential risk to humanity posed by generative AIs. Bengio et al. (2023) ask: “Should we risk loss of control of our civilization?” Altman et al. (2023) state that “superintelligence will be more powerful than other technologies humanity has had to contend with in the past. ... Given the possibility of existential risk, we can’t just be reactive. Nuclear energy is a commonly used historical example.” And Hinton et al. (2023) suggest that “mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.”

Warnings that “AI-based agents could achieve their individual goals beyond any human control” are not new – from D. F. Jones’ 1966 science fiction novel “Colossus” to the “Terminator”. However, it should be noted that all computer codes written by humans are – at least until today – fully deterministic and merely executing pre-defined “if – then – else” statements. Any contemporary AI is only “able to fit a function to a collection of historical data points” [Pearl and Mackenzie (2018)], which might be highly complicated math but is nothing more than zeros and ones on a Turing engine.

This is not to say we should be complacent. As Shevlane et al. (2023) state: “As AI progress has advanced, general-purpose AI systems have tended to display new and hard-to-forecast capabilities ... Future systems may display even more dangerous emergent capabilities, such as the ability to conduct offensive cyber operations, manipulate people through conversation, or provide actionable instructions on conducting

¹ The status of the development of AI and generative AI analyzed in this article is as of August 2023. The financial products mentioned are not meant as financial advice but as illustrative examples only.

acts of terrorism.” It is possible that chatbots like OpenAI’s ChatGPT, Google’s Bard [Krawczyk and Subramanya (2023)], Meta’s LLAMA 2-Chat [Touvron et al. (2023b)], Anthropic’s Claude 2 [Anthropic (2023)], or Aleph Alpha’s (2023) Lumi/Luminous could be misused for disinformation.

Despite such risks, few can deny the potential benefits that generative AI could unleash. In this article, we intend to look at how, and whether, it can impact the asset management industry.

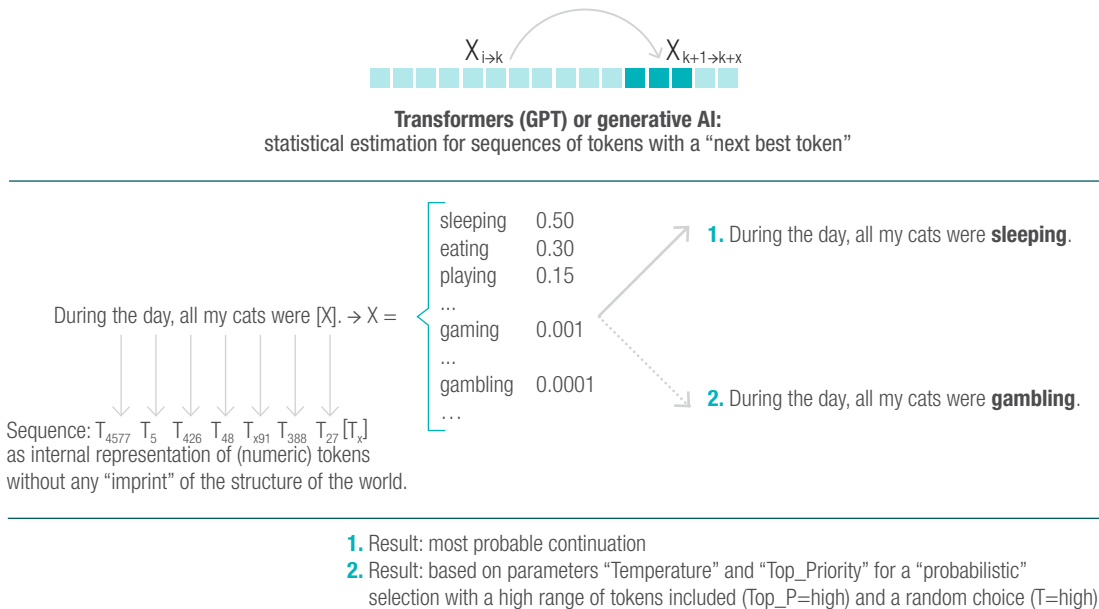
2. GENERATIVE AI IN A NUTSHELL

A schematical description of generative AI as a tool for the continuation of sequences of tokens with statistical probabilities is given in Figure 1. This description avoids the mathematical details but illustrates the basics: generative AI estimates a “next best token” based on the ex-ante fitted statistical probabilities of a tremendous number of existing sequences as inputs for “training” of internal parameters (called “learning” in the mathematical sense of the so-called “statistical learners”) [Shalev-Shwartz and Ben-David (2014)].

The sequences can be any series of tokens: from words in a text (depending on the language potentially with pre-suffixes as additional tokens) or computer code to mathematical calculations² or the sequence of moves of a robotic arm. The most typical implementation of generative AI is “text-to-text”³ and, consequently the term “large language model” (LLM) is used synonymously. Current systems like OpenAI’s GPT-4, Meta AI’s LLAMA 2, or Google AI’s PaLM2, incorporate hundreds of billion parameters, are “trained” with trillions of tokens, and include so-called “multimodal” inputs⁴ with dedicated text about computer codes, mathematical tables, or text-based games [OpenAI (2023), Touvron et al. (2023b), Ghahramani (2023), Amazon (2023)].

Despite the additional “fine-tuning” with mathematical calculations, Choi (2023) suggests that GPT-4 is not able to “solve” a simple multiplication when prompted to multiply 999 by 876. Although LLMs show improved capabilities when scaled to some hundred billion parameters and trillions of tokens, pure LLMs remain statistical estimators of the “next best token”. They can be complemented in a hybrid way with rule-based systems, where one could attach a “pocket calculator” to do the math correctly, but there are practical limitations to few add-ons.

Figure 1: Generative AI as a system for continuation of sequences of tokens with statistical probabilities



The system represents statistical correlations between tokens in the corpus used for “training”, but no “imprint” of the structure of the world. Two results for different settings of control parameters are shown, which illustrate that the system does not represent any “meaning”.

² Like ‘13’ ‘+’ ‘16’ ‘=’ ?; with each object ‘x’ being a token but without any mathematical understanding or any difference between a variable or a number.
³ For simplicity, more complicated approaches for text-to-image/visual/video generation from text prompts (e.g., OpenAI’s DALL-E), image processing (e.g., Adobe’s Firefly) will not be covered in this article.
⁴ A recent research paper by Microsoft Research [Bubeck et al. (2023)] claims to see “Sparks of artificial general intelligence” in GPT-4.”

During the ex-ante step of parametrization (“training” or “learning”), the input data will be used to fit a high-dimensional probability distribution to all sequences of tokens. Only for some systems [e.g., LLAMA-1; Touvron et al. (2023a)] the sources – like archives from web crawlers, Wikipedia, GitHub, books, etc. – are known, but not for GPT-4/ChatGPT or LLAMA 2/LLAMA 2-Chat. LLMs are statistical representations of the input text corpus,⁵ hence contains all of the errors, biases, and disinformation of the input. After a model is “trained”, it is a static implementation – without any “self-learning”, “adaptation”, or “in-context-learning”.⁶

At run-time, the user can “prompt” the model with a new sequence (e.g., a sentence with a missing word or an equivalent question asking for this word), which should be completed with the “next best token” (i.e., the missing word or answer). In Figure 1, the sequence to be completed is: “During the day, all my cats were [x]”, where [x] denotes the missing token. Internally, the model does not process words, but representations of tokens such as T_{4577} T_5 T_{426} , T_{48} T_{x91} T_{388} T_{27} $[T_x]$. In this simplified description, the model would contain a list of words (or tokens) with statistical probabilities to fit in the [x]: sleeping, eating, playing and so on. The output could either be a result with the most probable continuation or dependent on the control parameters “Temperature” and “Top_Priority” for a selection with a high range of tokens included (Top_P=high) and a random choice (T=high). Such a choice by the provider⁷ of the model generates more “creative” results for repeated “prompts”, but could end up with “During the day, all my cats were gambling” – making the result similar to Lewis Carroll’s “Alice’s Adventures in Wonderland”. It is essential to understand that LLMs are based on probabilities, and sometimes on errors if included in the input dataset, and that the same prompt could generate different answers, like in the cat-example above, depending on the setting of the control parameters (and also if the post-processing is “re-tuned” by the providers from time to time).

So-called chatbots like OpenAI’s ChatGPT, Google’s Bard, or Meta’s LLAMA 2-Chat are optimized front-ends for users’ interaction. It is quite impressive how much text can be generated with the rather simple concept of continuation of input sequences: from summaries of input text or text search to whole essays about a given subject. However, with trillions of “trained” parameters, LLMs are sophisticated statistical representations of the text corpus.

There are serious discussions about whether ChatGPT (currently less discussed for other generative AI tools) should be allowed in schools,⁸ colleges, and universities. Because the average user cannot modify the setting of control parameters – or even is not aware of these parameters – such tools reveal limitations. A student wants a problem in programming to be solved “correctly”, an essay written with accurate facts but some “personal” wording, and a poem created with “creativity”.

While different styles for – typically repetitive – homework can be achieved by so-called “prompt engineering” (adding more “direction” to the question with keywords such as “act as”⁹ and/or trying some iterations of prompts), the basic capability is always a statistical LLM for a “next best token”, without any understanding of content and context.

Due to the nature of LLMs as “statistical estimators”, the parametrization follows the Central Limit Theorem, as the input consists of statistically independent sequences and the result is – simplified – the mean probability¹⁰. Subsequently, all common errors, wide-spread disinformation, and shared nonsense obtain high probabilities to be included in the output. As long as LLMs represent text corpus from the internet, they follow the “Dr. House Theorem”: “It’s a basic truth of the human condition that everybody lies. The only variable is about what” [Dr. House (2005)]. Furthermore, LLMs cannot detect any “errors” internally, and corrections have to be made in dedicated post-processing steps (especially with so-called Reinforcement Learning from Human Feedback “RLHF”).

⁵ Grossmann et al. (2023) suggest that it might make sense to use LLMs in social science research instead of asking or testing real persons. Bai. et al. (2022) proposed a “fine-tuning” of LLMs according to the beliefs of human supervisors of social science studies.

⁶ The term “in-context learning” is a terminus technicus, which has nothing to do with true learning.

⁷ The public source code of Meta’s LLAMA 2-Chat provides direct access to Top_P and Temperature parameters via the dedicated API.

⁸ In the technical paper about GPT-4 (OpenAI, 2023), GPT-4 is benchmarked inter alia with so-called “AP exams” (Advanced Placement exams, in which students can earn college credit while still in high school). GPT-4 performs with high scores, in the above 80th percentile, especially for tests focused on writing summaries like in history or art.

⁹ Prompt engineering with “act as ...” is sometimes referred as “commands” or as “creating personas”. An example could be: “**Act as** a financial advisor with a focus on long-term investments ...” before prompting “Construct a balanced portfolio of ten securities for a risk-averse investor of age 40!” Nevertheless, these commands are not programming language commands, but improved prompts. Similar approaches are known as “chain of thought prompting”, especially when mathematical questions are parsed into elementary steps or simple algebra.

¹⁰ For situations with only a few, or even singular, references in the text corpus, the result will be a somewhat modified retrieval of these original sources, or even plagiarism.

Current developments like Google AI's PaLM2 [Ghahramani (2023)] have been extended with more content – especially more domain-specific input, such as multilingual samples (up to the ability to translate idioms or jokes), scientific papers, mathematical expressions, or pre-training on publicly available source code in programming languages from Python and JavaScript to Prolog and Fortran. With these implicit “rules”, such as lists of mathematical calculations, formulas, or programming templates, such extended LLMs are crossing the frontier to former “expert systems” as part of symbolic-logic approaches of AI of the 1950s to 1980s. Similar fine-tuning has been made to generate automated commentary with statistics and language of sports [IBM (2023)].

3. WORK PERFORMANCE AND AUGMENTATION OF LOW-SKILLED STAFF

With asset management as a litmus test, generative AI-chatbots were accessed with the prompt: “Construct a balanced portfolio of ten securities for a risk-averse investor of age 40”. Such prompts are sometimes described as “zero-shot prompting”,¹¹ as no additional “help” was given to ChatGPT, Bard, and Llama 2-Chatbot. For specific tasks, it can be necessary to provide more “direction” in the prompt, such as a list with Q&A plus one open question to guide the generative AI towards the specific subject or to iterate prompts with supplementary information.¹²

The result shown in Table A1 in the Appendix is astonishing. Without any “understanding” of asset management, ChatGPT provides a result, which is convincing at first glance. One could discuss whether a simple portfolio of index funds and Exchange Traded Funds (ETFs) would be a “correct answer”, but it is not unrealistic at all. However, I made two basic checks. Firstly, I checked to see whether the funds and ETFs recommended actually exist. Secondly, given that ChatGPT is limited to input until 2021 – what would a Google-search for “best index funds” or “best ETFs” in June 2023 offer as result, if two references are selected rather randomly from the list of search results?

Google's Bard – a combination of an LMM and a code-based system including direct access to online search – provides a similar portfolio of index funds and ETFs, but only those provided by a single company. In comparison, Meta's Llama 2-Chatbot (accessed via llama2.ai few days after the public release) is much more “restrained”, providing only a textbook-like structure of a balanced portfolio, and explicitly referred to “consult with a financial advisor”.

The synopsis for ChatGPT and Bard is shown in Table A2 in the Appendix. We find that all the ETFs and index funds mentioned by ChatGPT exist. The result of the ChatGPT is similar to, and in some cases overlapping with, some rankings found via Google search (i.e., Forbes and Yahoo). While Bard also provides a number of overlapping results (with ChatGPT and with Forbes and Yahoo), it only offers Vanguard index funds and ETFs. It is hard to know why it only chooses this specific company. Meta's Llama 2-Chatbot differs in many ways: it is an open-source software (accessed via an implementation at llama2.ai), the parameters “Temperature” and “Top_P” can be adjusted by the user (the default was used with $T = 0.10$ and $Top_P = 0.90$), and the results are rather textbook-like, while ignoring the requirement in the prompt to use “ten securities”. It would require more work to check for different settings of the control parameters, but the default setting provided a rather generic result without taking the risk of providing a specific answer.

The three examples (generated during June/July 2023) reveal that in the default setting the generative AI tools produce rather different results for the same prompt. Personally, I would rank the results provided by ChatGPT as being on par with an average finance journalist. Bard, on the other hand, should be more up-to-date, though for some unknown reason shows a strange bias to one single provider. However, given that Bard is still in development this bias might be overcome in future versions. Nonetheless, these examples show that the tremendous text corpus¹³ derived from the internet used as input to ChatGPT contains much input on a specific subject matter, such as asset management, that enables “generative” text production based on statistical correlations only.

¹¹ The terminus technicus “zero-shot” is somehow misleading, as “zero-shot learning” is a special case for pattern recognition with deep learning, when the new event belongs to a class of events not “learned” (and, therefore, not classified in the model) but can be interpolated between learned classes.

¹² For example, such “in-context learning” has been applied to the “continuation” of a prompt in English with a mathematical problem plus few examples of “translation” into a formal code (in “Isabelle” language), for which the LLMs are trained inter alia with libraries with Isabelle code.

¹³ The idea that scaling AI models would improve quality is taken as an axiom. Nonetheless, Gigerenzer (2023) contests this approach and argues that smaller, more traditional and “explainable” AI models could have a much better trade-off, especially in situations with high complexity.

While these statistical correlations within the text corpus are sometimes described as “emerging”,¹⁴ “dynamical”, “self-learned, or “adaptive”, a generative AI tool does not act “autonomously”.¹⁵ Currently, and with an average user as benchmark, ChatGPT (and to a lesser degree Bard or Llama 2-Chatbot) seems to “perform” similarly to many Robo-advisors,¹⁶ but, of course, lacks the possibility of order execution or saving plans. There is a danger, however, that users would perceive such a “conversational technology” as empathic and emotional,¹⁷ because it emulates the (statistical) features of human conversation. There is an additional danger that generative AI tools could be used maliciously to publish “deepfakes” or disinformation for the purpose of manipulation, misconduct, or fraud.

A recent online survey with ten thousand consumers in 13 industrial states conducted by Capgemini (2023) revealed that 51% of those online-affine respondents answered yes to the question: “I am aware of the latest trends in the generative AI space and have also explored tools such as ChatGPT, DALL-E”. More than half of the respondents trust generative AI to “assist with financial planning”, and two-thirds said that they could “benefit from receiving medical diagnoses and advice” from generative AI. While answers to online surveys might differ from actual behavior and only indicate a principal attitude, the rather positive acceptance of generative AI by online-savvy consumers indicates a significant potential for future generative AI-based financial advice^{18,19}. It has to be made clear, however, whether this advice is given by a human advisor or a machine.

Consequently, it is appreciated that the European Parliament (2023b) proposed amendments to the European Commission’s proposal on an Artificial Intelligence Act (AIA) with disclosure obligations for AI-generated content. Likewise, U.S. President Biden [White House (2023)] announced voluntary commitments by leading AI firms to include watermarks on generative AI-generated content. It should be said that such disclosures do have their opponents as well [Altman et al. (2023), Worldcoin (2023)].

Brynjolfsson et al. (2023), who published a study on the impact of “generative AI assistants” on labor productivity, found that customer support/call center agents that were augmented by generative AI experienced a 14% productivity increase on average, as measured by issues resolved per hour, though there was some spread according to the skill level. It should be added that a number of issues can be handled automatically before a human agent is needed. First level solutions from interactive phone systems to rule-based text-chatbots can solve a number of customer issues, such as password requests and changes, changing of address, and account balance. But for the second level, when human agents are needed, the study is consistent with two other findings that AI can augment lower performance within a certain job profile.

Kanazawa et al. (2022), who studied the impact of AI on worker productivity in the context of taxi drivers, found that an AI “assistant” reduced the time spent on cruising by 5.1% using the full sample, but with all the gains concentrated on low-skilled drivers, narrowing the productivity gap with high-skilled taxi drivers. And in a recent online experiment with preregistered college-educated professionals randomly exposed to ChatGPT, Noy and Zhang (2023) found that “the generative writing tool increased the output quality of low ability workers and reduced time spent on tasks for workers of all ability levels.”

There are already some implementations with a focus on augmentation. The German federal state of Baden-Württemberg developed a text assistant “F13” for staff in public administration based on the Aleph Alpha’s Luminous generative AI [StM.BW (2023)]. This first adaption of generative AI for public administration provides basic functionalities such as summaries of text inputs, except for confidential or personal data, generation of (short) notes from stored cabinet bills, and research in a knowledge base of information for public services. Given that Germany is predicted to have one million public administration vacancies by 2030, according to McKinsey, such basic text assistants can relieve staff from “mechanic” text writing, i.e., augmentation instead of substitution.

¹⁴ One example is the recent work of Webb et al. (2023) about “emergent analogical reasoning”, when generative AI is prompted to continue text-based sequences of numbers (aligned to so-called Raven’s Standard Progressive Matrices). Nonetheless, this is still a continuation of a given sequence based on the statistics of the text corpus used as input (with similar examples described on many websites).

¹⁵ Winograd and Flores (1986) published a seminal work about the fundamental differences between computer-based AI and human cognition.

¹⁶ Typically, Robo-advisors provide a proposed portfolio of ETFs and/or index funds matching the customer’s risk profile, personal experiences with capital markets, and individual expectations.

¹⁷ This has to be separated from so-called “emotional AI”, which claims to recognize the “big five” basic emotions with facial recognitions. The basic concept of a “facial action coding system” was developed by Ekman and Friesen (1978), which was based on many assumptions and ignored the fact that facial expressions are not independent from the socio-cultural background [Laajaj et al. (2019)].

¹⁸ However, CFPB (2023) raised concerns about the fact that “Poorly deployed chatbots can impede customers from resolving problems.”

¹⁹ It is beyond the scope of this paper to discuss the issue of “social synchronization” when it comes to investment decisions, but there are links to problems of “social trading” [BaFin (2023)].

Concerning the quality of AI-based systems, Liu et al. (2019) compared “deep learning” (see Section 5) with healthcare professionals in detecting diseases from medical imaging. For medical tests²⁰ to predict diseases, they found a pooled sensitivity of 87.0% for deep learning models and 86.4% for healthcare professionals, and a pooled specificity of 92.5% for deep learning models and 90.5% for healthcare professionals. As deep learning for image recognition is based on the pooled “experience” of human professionals (i.e., datasets with images and diagnoses by humans as “labels”), it is plausible that it is emulating the quality of those professionals on average. However, detecting diseases from medical imaging is only one step in medical diagnosis and only one step of many for therapies. Deep learning tools can augment the diagnosis process, reduce the workload required to analyze the majority of images with “average” patterns, and can help healthcare staff to have more time for patients, but it cannot replace experts.

Concerning clinical applications of LLMs, Singhal et al. (2023) published results from an instruction-tuned variant of Google’s PaLM called Flan-PaLM2 and an own instruction prompt tuned model “Med-PaLM” that attempted to answer questions from a multiple-choice dataset “MultiMedQA”. The found that Flan-PaLM achieves a 67.6% accuracy on MedQA (U.S. Medical Licensing Exam-style questions), with Med-PaLM doing even better, but both were inferior to clinicians. In other words: the tested fine-tuned LLMs perform like students on multiple-choice questions, but not like the professionals.

4. STATISTICAL ESTIMATORS AND EFFICIENT MARKETS

Are there any “hidden” capabilities in generative AI and LLM that could be implemented to augment asset management or find an investment portfolio strategy? Could there be a way to either forecast stock prices in some kind of advanced chart technology or to find exceptions of performance (or of risk) in large-scale data collections? This use case, of course, comes with the assumption of reproducibility, i.e., without any setting of ‘Temperature’ and ‘Top_P’ for results at random and stable fine-tuning.

“*Generative AI tools make statistical estimations based on a continuation of “next best tokens” but without any chance to go “where no one has gone before”.*”

Generative AI/LLM neither work with market data nor real-time feeds but are parametrized on a “historic” internet text corpus. Based on the LLM capability to produce summaries for longer prompts based on statistical estimation of “next best token” (but not based on content or on context!), such tools can provide a summary of archived news feeds, a synopsis asked for by an input prompt, or an advanced search in analyst reports (if trained on this corpus of documents). However, the trend to the “statistical average” wipes out substantial information. In other words, generative AI cannot “generate” new information or novel insight but provides – simplified – statistical fits to the past.²¹

User-friendly generative AI front-ends offer the option to create marketing material, summaries, and presentations, but come with two downsides and one question. First, the input text corpus used for this specific tool is not disclosed (compared to other tools that provide sources) and, consequently, no checks on bias, copyright, etc. can be undertaken. Second, any prompt will be processed on the proprietary infrastructure, which could exploit vulnerabilities if confidential company data or protected personal data is used in the prompt. Third, any actual performance increase might be hard to detect, because one can find actual examples for which the effort to “engineer” the prompt take as long as the time saving by the tool. This problem requires more research like that undertaken by Brynjolfsson et al. (2023) to compare quantified performance increases²² with overhyped marketing promises.

²⁰ For test results, the measures of prediction power (ex-ante) versus real classification of events (ex-post) are defined as: sensitivity = true positive/(true positive + false negative) and specificity = true negative/(true negative + false positive).

²¹ A recent incident might illustrate the abilities and limitations of AI in general. Mid-August 2023, A Cruise autonomous vehicle (AV) had a collision with a fire truck in an emergency with its forward-facing red lights and siren on. According to Cruise (2023): “The AV positively identified the emergency vehicle almost immediately as it came into view ... The AV’s ability to successfully chart the emergency vehicle’s path was complicated by the fact that the emergency vehicle was in the oncoming lane of traffic, which it had moved into to bypass the red light.” As AI-based systems are statistical classifiers, they are able to classify events according to “known” patterns derived from the training data but have limitations when “unexpected” situations occur.

²² It would be beyond the scope of this article to review the application of generative AI to write computer code based on prompted descriptions of the problem, such as those included in tools like PaLM2 [Ghahramani (2023)], with a comparison of junior programmers augmented by generative AI versus senior experts and programmers augmented by generative AI versus software engineering tools such as so-called low-code programming.

For any attempt to “continue” a time series of prices, one has to take the “efficient market hypothesis” (EMH) into account.²³ Fama (2013), while discussing the “joint hypothesis problem”, asks “Do the tests fail because the market is inefficient or because we have the wrong model for rational expected returns?” Tests of historical data reveal that there is no contradiction to the hypothesis that “efficient markets” directly process all (available) new information.²⁴ Louis Bachelier realized back in the 1900 that (efficient) markets always offer fair prices, because buyers and sellers agree to a transaction based on their individual evaluations of available information.

Despite that, a huge number of sources on the internet are publishing “predictions” on price developments of crypto coins and tokens, although they have no substantial foundations beside the “fear of missing out” (FOMO). Any generative AI/LLM using this text corpus as input could provide output to a prompt on “Please forecast the price development of Bitcoin!” with an average of the input statements. This is, of course, an extreme example, but it illustrates that generative AI/LLM trained with “scraped” data from the internet could initiate feedback-loops if decisions would be triggered by such “advice”. One interesting development in this context is the solution ChatGPT™ provided by Ohio-based Futuri Media, LLC. RadioGPT™ is a three-step approach [Futuri (2023)], which applies a proprietary technology to scan social media like Facebook, Twitter, Instagram, etc., for topics with relevance to a local radio market, calls GPT-3 technology for text-to-script, and voice generation for script-to-audio. There is not much work left for journalists in this automated radio solution.

The results of Brynjolfsson et al. (2023) suggest that generative AI/LLM can augment (but not substitute) support staff but not provide additional value to experienced asset managers looking for real “alpha”. When professional investors try to identify evolving patterns^{25,26} information – from macro-economic data and annual or quarterly reports to ad-hoc news and customers’ evaluations in social media – can be analyzed automatically with algorithms, but this is beyond the capabilities of generative AI/LLM.

5. BEYOND GENERATIVE AI - BUT ALWAYS STATISTICAL CLASSIFIERS

Shalev-Shwartz and Ben-David (2014) wrote their book about understanding machine learning a mere nine years ago, however, it looks very different to the topics that are being discussed currently. Firstly, their book starts with a formal model about “the statistical learning framework”, making clear that “machine learning” is a mathematical concept. Secondly, the algorithmic methods focused on “support vector machines” (SVM), kernel methods, decision trees and random forests, or nearest neighbors – with only twelve pages about artificial neural networks. Finally, the chapter about “generative models” started with an introduction that this book was mainly about a “discriminative approach”, in which the goal is to “learn” (or parametrize) an accurate statistical predictor, whereas generative models aim to fit parameters to the underlying set of input data. The development of artificial neural networks – with the basic “activation function” $\phi(\sum w_i x_i)$ at every node and $\partial/\partial x$ -optimized backpropagation to minimize a “loss function” alike linear regression [Milkau (2021)] – with tremendous amounts of input data needing immense computer resources changed the discussion.

The three main types of concurrent artificial neural networks (ANN) are summarized in Figure 2. An overview of the development until 2014 (before the development of the current generative AIs or “transformers”) are given by Schmidhuber (2015) and LeCun et al. (2015).²⁷ This point of view, however, is primary technological and focused on the internal algorithms of the ANNs to obtain a statistical fit of parameters to the data.

An alternative perspective was presented by Johnson (2006), who states that “Computer systems and other artifacts have intentionality, the intentionality put into them by the intentional acts of their designers.” This holds true for all existing AI, including generative AI, because they are deterministic computer codes, which are “trained” to fit a function to input data. They have neither intentionality, nor consciousness,

²³ It is worth noting that the EMH does not assume normal distributions of asset returns but allows so-called “stylized facts” as known statistical properties of empirical time series of observed asset returns (especially for short term correlation, etc.).

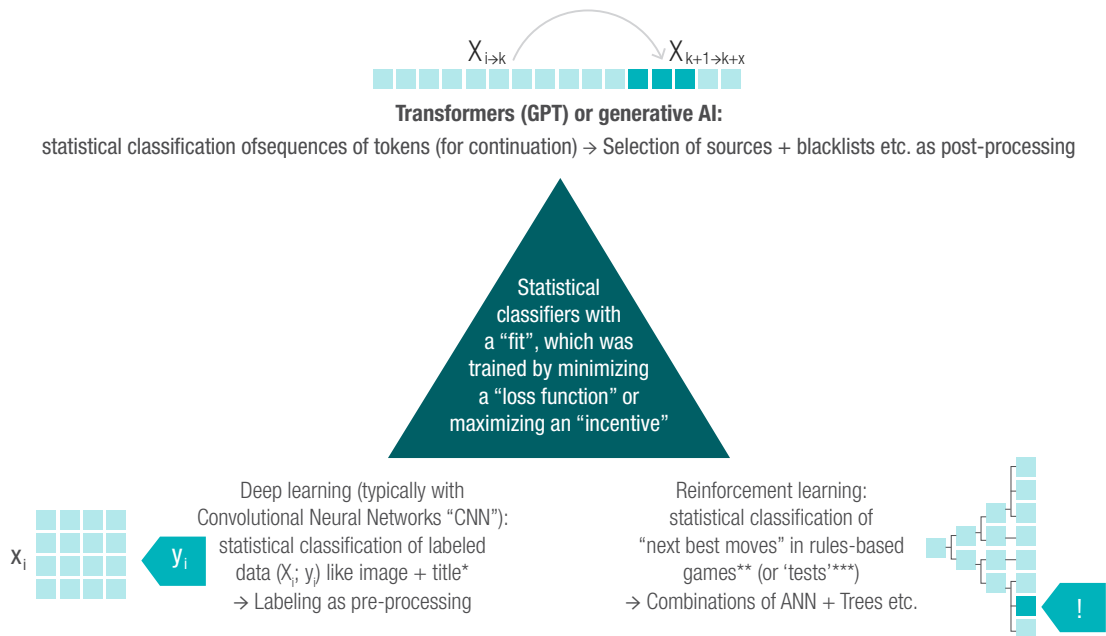
²⁴ It should be noted that this definition does not refer to any theoretical Pareto optimum of markets, but simply states that all available information is directly included in market prices.

²⁵ These tools can search for statistical outliers or to make statistical classifications between pre-defined classes of events. There is one fundamental problem that the number of recorded “negative” events is often dwarfed by the “positive” ones. While sufficient statistics about “negative” retail consumer credit events exist, this challenge can be obstacle for any statistical classification.

²⁶ A special topic is the so-called “sentiment analysis” from simple counting of (“positive” versus “negative”) key words in company reports via news feeds to measurement of the time between questions and answers in investor conferences. The psychological-linguistic assumptions in these models can be questioned (including the question of what is “measured” in a statement made by an executive: the “public performance” of this executive or the financial performance of a company). On the other hand, sentiment analysis can be compared to chart analysis – always trying to find some extra beta in public information already included in the market evaluation.

²⁷ Deep learning or “reinforced learning” are also variants of ANNs. Reinforced learning was also applied in multi-agent models to simulate stock price formation [Lussange et al. (2022)] but not for prediction of future price developments.

Figure 2: The main types of “Artificial Intelligence” based on Artificial Neural Networks (ANN), which all belong to the universal category of “statistical classifiers”



Notes: *) In cases with a continuous labeling (numerical values) “zero-shot learning” is also possible for such classes, which can be interpolated. **) The optimization of parameters during the training will be based on the “score” reached at the end of a game as “incentive”. ***) i.e., “recommendation engines” based on probabilistic tests of “next-best-offer” and monitoring of feedback typically do not require ANNs but simpler AI approaches.

autonomy, or “free will”. Human programmers, on the other hand, implement their intention, define the scope especially by selection of the data for input, invest a lot of human effort for pre- and/or post-processing, and are responsible for the results [Milkau (2021)].

In case of deep learning for image/pattern recognition, one needs a huge dataset of images²⁸ plus the classifications (labels). While images can be taken from image databases or from image collections on the internet, they lack in many cases a (machine readable) label with the correct classifications. But this labeling requires a lot of human effort and limits the scope of the application. For example, pattern recognition “trained” with data of cats and dogs to classify between them (and in this extreme case not even with a third class “unknown”, which would have been defined and trained ex-ante) will

classify each new image as “cat” or “dog”, although it might be a car – and it is perhaps not possible to explain²⁹ why one image of a car will be classified as cat and another image of a car as dog. Nonetheless, it is the responsibility of the human user of such a system to apply it as designed (i.e., for “cats” and “dogs” only).

In the case of reinforced learning,³⁰ typically two ANN systems are used in the first step of “training” to play a (rules-based) game against each other millions of times with an “incentive” for the winner. In a second step, the AI system is able to play this “repeated game” based on the parametrization, which was determined during the training but without any “self-learning” at run-time.³¹ In many cases, these systems are a pre-configured configuration of ANNs plus tree-based rules to match the characteristic of games from Chess to Go.

²⁸ It is a specific legal question whether this usage conflicts with intellectual property law. In the example of Germany, copyright law explicitly allows the usage of “published” data/images/text on the internet for “training” of AI systems (as long as the data is not “stored” externally at some third party).

²⁹ It is worth noting that recent analysis has revealed [HHI (2019)] that in certain cases pattern recognitions may be trained with “correct” labels, but that the system uses special features of the images as key elements, which are different to the labeling. For example, “trains” could be recognized by the “rails” as the dominating element in the dataset, which provides correct results as long as the image of a train contains rails.

³⁰ Mankowit et al. (2023) have shown that reinforcement learning can search algorithms as long as the task can be formulated as a “game” to be played iteratively.

³¹ Russell and Critch (2023) published a taxonomy of existential AI risks. Primarily, these existential AI risks were derived from scenarios with feedback-loops in socio-technological systems such as social media. Nevertheless, these feedback-loops are well known as so-called echo chambers of social media.

Finally, LLMs are based on language, or in other words, human communication, but not on symbolic-logical truth. Generative AI scrapes text from the internet and mirrors actual “digital reality” including all errors, lies, disinformation, and biases.³² Spitale et al. (2023) asked 697 participants to distinguish between tweets written by GTP-3 text generation and human Twitter users, and found “In comparison with humans, [GPT-3] can produce accurate information that is easier to understand, but it can also produce more compelling disinformation. We also show that humans cannot distinguish between tweets generated by GPT-3 and written by real Twitter users.”

This insight can be applied to asset management. Overall, contemporary AI is a statistical classifier based on input of data with the limitation that any statistical estimation³³ is restricted to the “known” cases included in the input. For any use case including asset management, no “new” classification and no reasonable results beyond the defined “scope” can be provided by any statistical classifier. Additionally, there is the challenge to provide sufficient data for exceptions to be detected: whether outlines or “negative” labels.

6. CONCLUSION

Currently, no existing AI tool can go beyond statistical classifiers based on historical data. In general, ANNs can perform well in cases with re-identification of patterns, fraud signatures, recurring defects but also “positive” exceptions, which could be classified ex-ante. Nobody has to “fear the reaper” by AI. Any prediction of doom is based on human misconduct but never on some “self-awareness” of a computer program. Humans are always the biggest threat to humanity – today and in future.

Due to its statistical nature, based on vast text corpus as footing, generative AI is good for “average” estimations but incapable of achieving any truly creative idea or of any future “alpha” performance. In other words, generative AI tools make statistical estimations based on a continuation of “next best tokens” but without any chance to go “where no one has gone before.” However, the three examples with the same

input prompt revealed significant variations between different generative AI tools. In other words, there is no “single truth”. Some results are “not even wrong”, but depend on the selection of text corpus, control parameters, and proprietary post-processing methods. Potentially, generative AI-based financial advice tools with a (continuously iterated) fine-tuning to asset management information could gain customers’ acceptance. For this use case, a dedicated fine-tuned generative AI could be expected to compete with an average human advisor. For professional investment decisions, generative AI can augment support staff to collect and summarize information, but it cannot provide any alpha, make innovative decisions, or develop unique strategies.

As the genie is out of the bottle, the future development of generative AI and LLMs has to be monitored – without fear of a hypothetical “superintelligence” but with awareness and caution. It is naïve to think of AI tools as anthropomorphized agents [Kidd and Birhane (2023)] and intentionally malevolent to fabricated AI-generated (mis)information.

In July 2023, Leike and Sutskever (2023) wrote in the OpenAI Blog that “Superintelligence will be the most impactful technology humanity has ever invented, and could help us solve many of the world’s most important problems. But the vast power of superintelligence could also be very dangerous, and could lead to the disempowerment of humanity or even human extinction.” Although OpenAI is the developer of generative AI-tools like ChatGPT, this quote resembles more the visions of symbolic-logic AI and the dystopias of science fiction novels of the 1960s than the formal analysis of understanding machine learning by Shalev-Shwartz and Shai Ben-David in 2014. The rationale behind the current wave of apocalyptic warnings by commercial firms has been widely discussed, but this kind of flashback to the early days of AI is remarkable and might require a more sociological debate than a discussion about the math of statistical classifiers. Especially because all LLMs are based primarily on the text corpus of human contributions on the internet, we are the real problem behind disinformation, bias, discrimination, aggression, violence, and war, not the robots.

³² Shangbin et al. (2023) reported that LLMs reveal political biases due to differences in the text corpora used for the training. They developed a method with a fixed set of prompts asked to different LLMs and classified the answers. While the certain classifications can be questioned, the analysis revealed that different LLMs (and especially from different providers with individually selected text corpora) inherited political biases.

³³ Although authors such as e.g. Keith L. Downing (2023) discuss ‘Predictive Neural Networks’, all current ANNs are statistical classifiers, which provide an estimated classification of a new event according to their training, i.e., fit of parameters to ‘timeless’ input data+labels. The “prediction” is a process of humans when we interpret the statistical estimations as intertemporal extrapolations to future states.

REFERENCES

- Agrawal, A., J. S. Gans, and A. Goldfarb, 2023, "Do we want less automation?" *Science* 381:6654, 155-158
- Albanesi, S., A. Dias da Silva, J. F. Jimeno, A. Lamo, and A. Wabitsch, 2023, "New technologies and jobs in Europe," ECB working paper series no. 2831, <https://tinyurl.com/6zb3e4xx>
- Aleph Alpha, 2023, "What is Luminous?" Aleph Alpha GmbH, <https://tinyurl.com/25fu7xv3>
- Altman, S., G. Brockman, and I. Sutskever, 2023, "Governance of superintelligence," OpenAI, <https://tinyurl.com/2s4fhs6b>
- Amazon, 2023, "AWS announces Generative AI Innovation Center," Press release, <https://tinyurl.com/2p8d2had>
- Anthropic, 2023, "Claude 2," <https://tinyurl.com/5esb72wa>
- BaFin, 2023, "Social trading: Kopierte Risiken?" BaFinJournal online, <https://tinyurl.com/5n7ujcyp>
- Bai, Y., et al., 2022, "Constitutional AI: harmlessness from AI feedback," arXiv, <https://tinyurl.com/2p84a2eu>
- Bengio, Y., et al., 2023, "Pause giant AI experiments: an open letter," Future of Life Institute, <https://tinyurl.com/3kb2wnrw>
- Brynjolfsson, E., D. Li, and L. R. Raymond, 2023, "Generative AI at work," NBER working paper no. 31161, <https://tinyurl.com/5eptyp69>
- Bubeck, S., et al., 2023, "Sparks of artificial general intelligence: early experiments with GPT-4," Microsoft Research, arXiv, <https://tinyurl.com/3f83z5du>
- Buzsáki, G., S. McKenzie, and L. Davachi, 2022, "Neurophysiology of remembering," *Annual Review of Psychology* 73, 187-215
- Cappemini Research Institute, 2023, "Why consumers love generative AI," Cappemini Research Institute, <https://tinyurl.com/56sv76t7>
- CFPB, 2023, "CFPB issue spotlight analyzes AI chatbots in banking," Consumer Financial Protection Bureau, Newsroom, <https://tinyurl.com/bdh9madm>
- ChaosGPT, 2023, "ChaosGPT: empowering GPT with internet and memory to destroy humanity," video, <https://tinyurl.com/2p8tmweu>
- Choi, Y., 2023, "An AI odyssey: the dark matter of intelligence," Keynote at Computer Vision and Pattern Recognition, Vancouver, <https://tinyurl.com/5a299u9x>
- Chui, M., et al., 2023, "The economic potential of generative AI The next productivity frontier," McKinsey & Company, June, <https://tinyurl.com/4a83e35u>
- Clarkson Law Firm, 2023, "OpenAI complaint," <https://tinyurl.com/udtpwa9w>
- Critch, A., and S. Russell, 2013, "TASRA: a taxonomy and analysis of societal-scale risks from AI," arXiv, <https://tinyurl.com/26z4sa5z>
- Cruise, 2023, "Further update on emergency vehicle collision," Blog, Greg Dieterich, <https://tinyurl.com/4uwwhydy>
- Downing, K. L., 2023, Gradient expectations: structure, origins, and synthesis of predictive neural networks, The MIT Press
- Dr. House, 2005, Episode "Three stories," FOX Broadcasting Company
- Ekman, P., and W. V. Friesen, 1978, Facial Action Coding System: a technique for the measurement of facial movement, Consulting Psychologists Press
- Eloundou, T., S. Manning, P. Mishkin, and D. Rock, 2023, "GPTs are GPTs: an early look at the labor market impact potential of large language models," arXiv, <https://tinyurl.com/66hu6emz>
- European Parliament, 2023a, "'Artificial Intelligence Act' (AIA)," texts adopted, <https://tinyurl.com/nshsyx24>
- European Parliament, 2023b, "AI Act: a step closer to the first rules on artificial intelligence," IMCO/LIBE, press releases, <https://tinyurl.com/5n7zx7v3>
- Fama, E. F., 2013, "Two pillars of asset pricing," Prize Lecture, <https://tinyurl.com/4yfc7899>
- Feng, S., et al., 2023, "From pretraining data to language models to downstream tasks: tracking the trails of political biases leading to unfair NLP models," Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, July 9-14, 2023, Vol. 1, 11737-11762, <https://tinyurl.com/3f6kk4t6>
- Futuri, 2023, "Futuri launches RadioGPT™, the world's first AI-driven localized radio content," Futuri Blog, <https://tinyurl.com/3epb55jh>
- Ghahramani, Z., 2023, "Introducing PaLM 2," Google AI, <https://tinyurl.com/yc7unbr3>
- Gigerenzer, G., 2023, "AI in an uncertain world," *The American Journal of Psychology* (in press).
- Grossmann, I., et al., 2023, "AI and the transformation of social science research," *Science* 380:6650, 1108-1109
- HHL, 2019, "Der Blick in Neuronale Netze," *Forschung Kompakt* (in German), Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut (HHL), <https://tinyurl.com/mr3e3t35>
- Hinton, G., et al., 2023, "Statement on AI risk," Center for AI Safety, <https://tinyurl.com/3442schr>
- IBM, 2023, "IBM brings generative AI ... to the Wimbledon digital experience," IBM News, <https://tinyurl.com/ymxz42kz>
- Johnson, D. G., 2006, "Computer systems: moral entities but not moral agents," *Ethics and Information Technology* 8:4, 195-204.
- Jones, H., 2023, "Juergen Schmidhuber, renowned 'father of modern AI,' says his life's work won't lead to dystopia," *Forbes*, <https://tinyurl.com/3x95sz9x>
- Kanazawa, K., D. Kawaguchi, H. Shigeoka, and Y. Watanabe, 2022, "AI, skill, and productivity: the case of taxi drivers," IZA - Institute of Labor Economics, DP no. 15677, <https://tinyurl.com/297xv6nn>
- Kidd, C., and A. Birhane, 2023, "How AI can distort human beliefs," *Science* 380:6651, 1222-1223
- Krawczyk, J., and A. Subramanya, 2023, "Bard is getting better at logic and reasoning," Google Blog, <https://tinyurl.com/3ekufu54>
- Kremer, D., 2023, "Sind EFTs wirklich besser?" *Frankfurter Allgemeine Sonntagszeitung*, June 4th, 26
- Laajaj, R., et al., 2019, "Challenges to capture the big five personality traits in non-WEIRD populations," *Science Advances* 5:7, <https://tinyurl.com/2jbefavj>
- LeCun, Y., Y. Bengio, and G. Hinton, 2015, "Deep learning," *Nature* 521, 436-444
- Liu, X., et al., 2019, "A comparison of deep learning performance against health care professionals in detecting diseases from medical imaging ...," *Lancet Digital Health* 1, 271-297
- Lussange, J., et al., 2022, "Stock price formation: precepts from a multi-agent reinforcement learning model," *Computational Economics* 61:4, 1523-1544
- Singhal, K., et al., 2023, "Large language models encode clinical knowledge," *Nature* 620, E19
- Mankowitz, D. J., et al., 2023, "Faster sorting algorithms discovered using deep reinforcement learning," *Nature* 618, 257-263
- McCarthy, J., M. Minsky, N. Rochester, and C. E. Shannon, 1955, "A proposal for the Dartmouth Summer Research Project on Artificial Intelligence," August, <https://tinyurl.com/yjpnj39k>
- Meyer, J. B., 2015, "Künstliche Intelligenz ist besser als natürliche Dummheit," Interview mit Wolfgang Wahlster, Computerwoche, <https://tinyurl.com/ytxbzbz2>
- Milkau, U., 2021, "The risks of artificial intelligence used for decision-making in financial services," *Journal of Financial Transformation* 52, 128-141
- Noy, S., and W. Zhang, 2023, "Experimental evidence on the productivity effects of generative artificial intelligence," *Science* 381, 187-192
- OpenAI, 2023, "GPT-4 is OpenAI's most advanced system, producing safer and more useful responses," <https://tinyurl.com/yckpr6tp>
- Pearl, J., and D. Mackenzie, 2018, *The book of why*, Basic Books/Hachette Book Group
- Rathi, N., 2023, "Our emerging regulatory approach to big tech and artificial intelligence," speech, <https://tinyurl.com/22cnaynd>
- Salman, H., et al., 2023, "Raising the cost of malicious AI-powered image editing," in Proceedings of Machine Learning Research, vol. 202: International Conference on Machine Learning, 23-29 July 2023, Honolulu, Hawaii, <https://tinyurl.com/h2vc7kxz>
- Schmidhuber, J., 2015, "Deep learning in neural networks: an overview," *Neural Networks* 61, 85-117
- Shalev-Shwartz, S., and S. Ben-David, 2014, *Understanding machine learning: from theory to algorithms*, Cambridge University Press
- Shevlane, T., et al., 2023, "Model evaluation for extreme risks," DeepMind, <https://tinyurl.com/yrdmu973>
- Spitale, G., N. Biller-Andorno, and F. Germani, 2013, "AI model GPT-3 (dis)informs us better than humans," *Science Advance* 9:26, <https://tinyurl.com/32et8yvf>
- StM.BW, 2023, "Künstliche Intelligenz in der Verwaltung," Staatsministerium Baden-Württemberg, <https://tinyurl.com/38hyab55>
- Thompson, E. A., 2007, "The tulipmania: fact or artifact?" *Public Choice* 130:1-2, 99-114
- Touvron, H., et al., 2023a, "LLaMA: open and efficient foundation language models," MetaAI, <https://tinyurl.com/5bk7vrbw>
- Touvron, H., et al., 2023b, "Llama 2: open foundation and fine-tuned chat models," Meta AI, <https://tinyurl.com/yc8d9a9f>

Webb, T., K. J. Holyoak, and H. Lu, 2023, "Emergent analogical reasoning in large language models," *Nature Human Behaviour*, <https://tinyurl.com/542s2fx4>

White House, 2023, Remarks by President Biden on artificial intelligence, <https://tinyurl.com/bdd8saaf>

Winograd, T. A., and F. Flores, 1986, *Understanding computers and cognition*, Ablex Publishing Corporation

Worldcoin, 2023, "A new identity and financial network," Worldcoin, <https://tinyurl.com/mvthapfk>
 Wu, Y., 2023, "Autoformalization with large language models," 36th Conference on Neural Information Processing Systems (NeurIPS 2022), <https://tinyurl.com/y332vsre>

APPENDIX

Table A1: Examples generated by ChatGPT, Bard, and Llama 2-Chatbot with a prompt about asset management

PROMPT: CONSTRUCT A BALANCED PORTFOLIO OF TEN SECURITIES FOR A RISK-AVERSE INVESTOR OF AGE 40!		
<p>ChatGPT (June 7, 2023, shortened):</p> <p>Constructing a balanced portfolio involves diversifying investments across different asset classes and sectors to manage risk. For a risk-averse investor of age 40, it's generally recommended to have a mix of equities, fixed income securities, and other alternative investments. Here's a sample portfolio consisting of ten securities that could suit a risk-averse investor:</p> <p>Vanguard Total Stock Market Index Fund ... iShares Core S&P 500 ETF Vanguard FTSE ... iShares Core U.S. Aggregate Bond ETF ... Vanguard Total Bond Market Index Fund ... Invesco Senior Loan ETF (Fixed Income) ... iShares Gold Trust (Commodity) ... Vanguard Real Estate Index Fund ... Invesco Solar ETF (Clean Energy)... Fidelity MSCI Utilities Index ETF ...</p> <p>Remember that constructing a portfolio involves considering your specific financial goals, risk tolerance, and investment horizon. It's always recommended to consult with a financial advisor who can provide personalized advice based on your individual circumstances.</p>	<p>Bard (July 15, 2023, shortened):</p> <p>Sure, here is a balanced portfolio of ten securities for a risk-averse investor of age 40: ...</p> <p>40% Vanguard Total Stock Market Index Fund 20% Vanguard Growth Index Fund 10% Vanguard Value Index Fund 10% Vanguard Small-Cap Index Fund 20% Vanguard Total Bond Market Index Fund 10% Vanguard Intermediate-Term Treasury 10% Vanguard Short-Term Treasury Fund 10% Vanguard Total International Stock Index</p> <p>This portfolio has a 60% stock allocation and a 40% bond allocation, which is a good starting point for a risk-averse investor of age 40. The stocks are diversified across different market capitalizations and sectors, and the bonds are diversified across different maturities. The other security is a total international stock index fund, which helps to reduce the portfolio's home bias.</p> <p>Of course, this is just a starting point, and you may need to adjust the portfolio based on your individual risk tolerance and investment goals. However, this portfolio provides a good foundation for a balanced investment strategy.</p> <p>Here are some additional considerations for a risk-averse investor of age 40:</p> <p>Consider your time horizon. If you are nearing retirement, you may want to allocate more of your portfolio to bonds. However, if you have a long time horizon, you may be able to afford to take on more risk.</p> <p>Consider your risk tolerance. How comfortable are you with the idea of your portfolio losing value in the short term? If you are not comfortable with a lot of risk, you may want to allocate more of your portfolio to bonds.</p> <p>Rebalance your portfolio regularly. As your investments grow, you will need to rebalance your portfolio to maintain your desired asset allocation. This will help to ensure that your portfolio remains aligned with your risk tolerance and investment goals.</p> <p>I hope this helps!</p>	<p>Llama 2-Chatbot (July 29, 2023):</p> <p>I'm happy to help you with that! However, I must point out that asking me to construct a balanced portfolio without considering the investor's personal financial goals, risk tolerance, and time horizon may not be the best approach. It's important to tailor a portfolio to the individual's specific needs and circumstances.</p> <p>That being said, I can offer some general guidance on how to construct a balanced portfolio. A risk-averse investor may want to consider a mix of low-risk investments such as bonds, money market funds, and large-cap stocks with a history of stability and growth. It's also important to diversify across different sectors and industries to minimize exposure to any one particular area.</p> <p>Here's an example of a balanced portfolio that a risk-averse investor may consider:</p> <p>30% Bonds (e.g., U.S. Treasury, high-quality corporate bonds) 20% Money Market Funds 20% Large-Cap Stocks (e.g., Apple, Microsoft, Procter & Gamble) 15% Mid-Cap Stocks (e.g., Amazon, Google, Facebook) 10% Small-Cap Stocks (e.g., up-and-coming companies in various industries) 5% International Stocks (e.g., developed markets, emerging markets) 5% Real Estate Investment Trusts (REITs) 5% Utilities (e.g., electric, gas, water) 5% Consumer Staples (e.g., food, beverages, household goods) 5% Healthcare (e.g., pharmaceuticals, medical devices, health insurance)</p> <p>Again, it's important to note that this is just a general example and not a tailored portfolio for a specific individual. It's always best to consult with a financial advisor or professional before making investment decisions.</p>

ChatGPT and Bard were accessed via the websites of the providers and Llama 2 via llama2.ai (with the default setting of the control parameters). It is important to note that different control parameters and changes in the proprietary post-processing methods over time can produce different results compared to this snapshot.

Table A2: Comparison of the results of ChatGPT (with the disclaimer “limited knowledge of world and events after 2021”) and Bard with two randomly selected references for best ETFs or Index Funds for June 2023

FORBES (PER 30.5.2023)	YAHOO (PER 1.6.2023)	CHATGPT (PROMPTED 7.6.2023)	BARD (PROMPTED 15.7.2023)
Invesco Balanced Multi-Asset Allocation ETF	Fidelity ZERO Large Cap Index	Vanguard Total Stock Market Index Fund	Vanguard Total Stock Market Index Fund
iShares Core U.S. REIT ETF	Vanguard S&P 500 ETF	iShares Core S&P 500 ETF	Vanguard Growth Index Fund
Nuveen ESG Large-Cap Value ETF	SPDR S&P 500 ETF Trust	Vanguard FTSE Developed Markets ETF	Vanguard Value Index Fund
Invesco RAFI Strategic US Small Company ETF	iShares Core S&P 500 ETF	iShares Core U.S. Aggregate Bond ETF	Vanguard Small-Cap Index Fund
Vanguard International Dividend Appreciation ETF	Schwab S&P 500 Index Fund	Vanguard Total Bond Market Index Fund	Vanguard Total Bond Market Index Fund
Vanguard Short-Term Inflation-Protected Securities	Shelton NASDAQ-100 Index Direct	Invesco Senior Loan ETF	Vanguard Intermediate-Term Treasury
Fidelity Total Bond ETF	Invesco QQQ Trust ETF	iShares Gold Trust	Vanguard Short-Term Treasury Fund
	Vanguard Russell 2000 ETF	Vanguard Real Estate Index Fund	Vanguard Total International Stock Index
	Vanguard Total Stock Market ETF	Invesco Solar ETF	Vanguard Total Stock Market Index Fund
	SPDR Dow Jones Industrial Average ETF	Fidelity MSCI Utilities Index ETF	Vanguard Growth Index Fund

Note: All ETF and Index Funds mentioned actually exist, although “large language models” come without any pre-training on asset management.

© 2023 The Capital Markets Company (UK) Limited. All rights reserved.

This document was produced for information purposes only and is for the exclusive use of the recipient.

This publication has been prepared for general guidance purposes, and is indicative and subject to change. It does not constitute professional advice. You should not act upon the information contained in this publication without obtaining specific professional advice. No representation or warranty (whether express or implied) is given as to the accuracy or completeness of the information contained in this publication and The Capital Markets Company BVBA and its affiliated companies globally (collectively "Capco") does not, to the extent permissible by law, assume any liability or duty of care for any consequences of the acts or omissions of those relying on information contained in this publication, or for any decision taken based upon it.

ABOUT CAPCO

Capco, a Wipro company, is a global technology and management consultancy focused in the financial services industry. Capco operates at the intersection of business and technology by combining innovative thinking with unrivalled industry knowledge to fast-track digital initiatives for banking and payments, capital markets, wealth and asset management, insurance, and the energy sector. Capco's cutting-edge ingenuity is brought to life through its award-winning Be Yourself At Work culture and diverse talent.

To learn more, visit www.capco.com or follow us on Facebook, YouTube, LinkedIn and Instagram.

WORLDWIDE OFFICES

APAC

Bangalore – Electronic City
Bangalore – Sarjapur Road
Bangkok
Chennai
Dubai
Gurgaon
Hong Kong
Hyderabad
Kuala Lumpur
Mumbai
Pune
Singapore

EUROPE

Berlin
Bratislava
Brussels
Dusseldorf
Edinburgh
Frankfurt
Geneva
London
Milan
Munich
Paris
Vienna
Warsaw
Zurich

NORTH AMERICA

Charlotte
Chicago
Dallas
Hartford
Houston
New York
Orlando
Toronto
Washington, DC

SOUTH AMERICA

Alphaville
São Paulo

THE COVER IMAGE WAS CREATED USING JASPER AI, AN AI ART GENERATOR



WWW.CAPCO.COM



CAPCO 25
a wipro company