RockCreek

Integrating Generative AI Into Investment Processes

Boosting Productivity & Gaining Competitive Advantage

Pooya Azadi, Sudhir Krishnamurthi

October 2024

1133 Connecticut Avenue N.W. | Washington, D.C. 20036 | Telephone: (202) 331-3400 | Facsimile: (202) 331-3420

KEY TAKEAWAYS

GenAI Provides Insights & Automates Tasks

1

GenAI solutions offer investment firms opportunities to boost productivity by automating routine tasks. They can also help firms gain a competitive advantage by generating insights, provided the firm can identify the right combinations of input data and processing methods that lead to outputs with high predictive power.

Data, Expertise & Iteration

2

Successful GenAI integration in investment workflows depends on several factors that include robust data preprocessing pipelines, contributions from subject matter experts during the ideation and testing phases, and the ability of the team to improve model performance based on qualitative feedback from users.

GenAI vs Conventional Machine Learning

3

While the general ideas and methodology underpinning the development of large language models (LLMs) share similarities with other types of machine learning (ML) models, the solutions enabled by GenAI exhibit distinct characteristics from traditional ML models. GenAI models, unlike conventional ML models, can operate with smaller unlabeled datasets. However, evaluating GenAI solutions is more challenging due to their open-ended interactions, necessitating iterative human feedback. Overall, the development of solutions based on traditional ML models rely more heavily on the technical data science skills.

Mitigating Risk is Critical

4

Beyond hallucination and algorithmic bias, which are broadly recognized as risks associated with the use of GenAI models, the proliferation of these tools within organizations could create an exaggerated notion about the depth, breadth, and objectivity of the insights. It has also been observed that the use of GenAI in groups can adversely affect collective creativity as it leads to less diverse ideas.

Focus on Creating Business Value

5

Although many business leaders agree that GenAI will disrupt their industries in the not-too-distant future, realizing the potential of GenAI has proven to be much harder than it appears. While GenAI has exciting potential, investment firms should take a pragmatic approach to implementation, focusing on creating demonstrable value through specific use cases rather than pursuing overly broad goals, which often lead to project failures.

INTRODUCTION

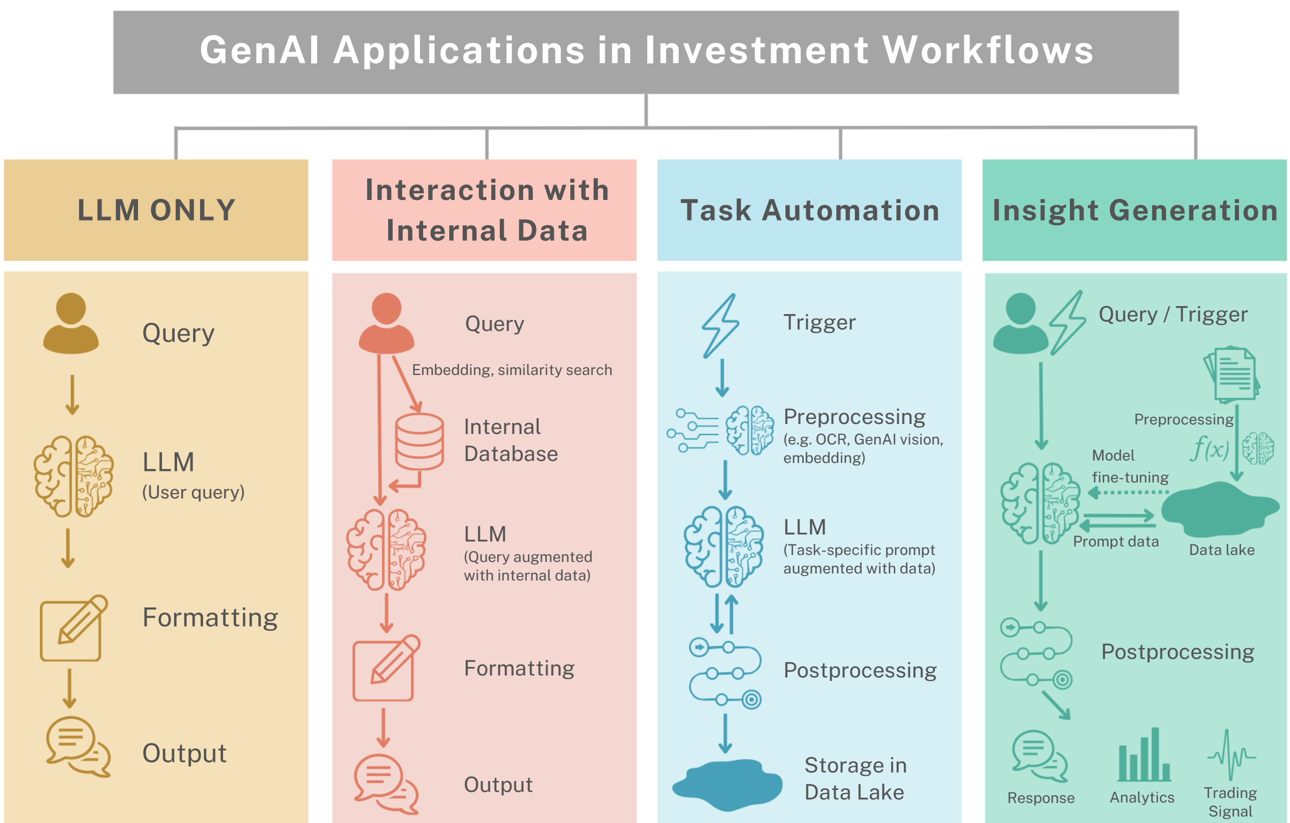
This paper provides a brief overview of how generative artificial intelligence (GenAI) can be effectively leveraged by investment firms to improve productivity and gain a competitive edge. The first part of the paper explores the broad categories of tasks and applications in investment activities that can readily benefit from utilizing LLM (large language models). The second part then provides examples of LLM-based applications that we have incorporated in our investment and operations workflows, as well as important lessons learned during the process.

Advances in the development of large language models have ushered in a sense of euphoria about AI and resulted in a new wave of AI applications adopted by companies across the industries. Writing article summaries or product descriptions, suggesting software code, generating natural language reports from structured and unstructured data, translating documents, performing sentiment analysis, and serving as the backbone of virtual assistants and chatbots are some examples of how organizations are currently leveraging GenAI to create business value. The most common applications of GenAI thus far have been focused on key functional areas such as sales and marketing (for content generation), product and R&D (for design ideas and testing), as well as IT and software development (for help desk and code generation) [1]. While many business leaders believe that GenAI will disrupt their industries in the not-too-distant future, realizing the sheer potential of this new technology has proven to be much harder than it appears at first glance [2]. The rest of this paper discusses some promising applications of GenAI for investment firms and analyzes their impact on organizations in terms of gains in productivity and/or competitive advantages.

GENERATIVE AI IN THE FINANCIAL INDUSTRY

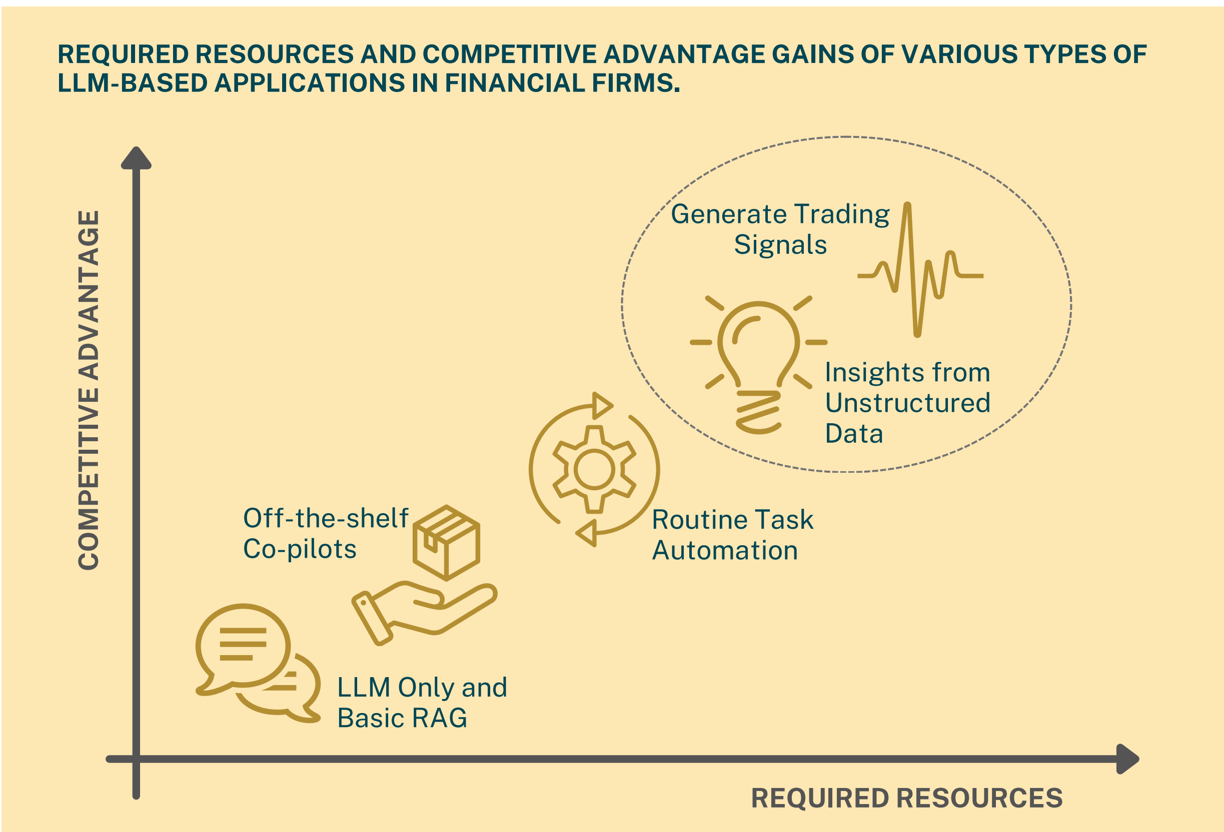
Besides robotic process automation, virtual agents and natural language processing are by far the most widely used AI capabilities in the financial industry [4]. These uses align well with the core strength of LLMs—namely their ability to process everyday language and to reason. To effectively leverage these capabilities in the financial industry, LLM models have to be fed with high-quality input data and optimized prompt(s); and in some cases, external modules and agents to perform complex calculations and tasks.

According to a recent survey, the financial industry currently spends an average of 7% of its digital budget on GenAI solutions [1], with a small fraction (e.g., 10–15% [3]) of this budget allocated for actual model costs. Most of the remaining budget is spent on the development of components for a reliable production-grade solution and on staff costs.



Benefits and Required Resources

The potential impact of GenAI solutions in financial firms can vary widely, ranging from a limited productivity boost at the individual level (that may not translate proportionally to higher productivity at the firm level without organizational restructuring) to gaining a substantial competitive advantage with a material impact on the firm's performance. The figure below provides a high-level conceptual comparison of the required resources and their potential impact for different levels of GenAI adaptation within financial firms.



LLM Only, Basic RAG, and Co-pilots

Creating private access to a fully-managed LLM on the cloud (e.g., AWS Bedrock) on a firm's proprietary platform and purchasing off-the-shelf product subscriptions (e.g., GitHub Copilot and Microsoft 365 Copilot) to assist employees completion of ad-hoc tasks are often the first steps firms take to incorporate new AI capabilities in their workflow. Despite the significant potential to boost productivity at the individual level, the integration of these off-the-shelf tools may not bring about substantial competitive advantages for the firm considering the commoditized nature of these tools and the low technical barriers for their implementation.

Task Automation

Creating tools to automate routine tasks specific to the firm constitutes yet another promising use case for GenAI in the financial industry. Extraction of key-value pairs from documents (e.g. exposures and performance) that the firm receives from outside sources (or generates through systematic web crawling) and transforming the extracted values into a target format to be uploaded into a structured database is one example of task automation. These applications can run in batch mode (e.g. extracting a specific field from a large number of documents in the company's data lake for research purposes), on a regular basis (e.g., periodically crawling a website to check for specific updates) or triggered by a pre-defined event (e.g., receipt of an email with file attachments from specific sources). The following sections provide other examples of task automation solutions that we have implemented at RockCreek.

The expected return on investment (ROI) of task automation projects depends on the resources required to build the task automation pipeline (and its future maintenance) and the total amount of time and financial resources saved through automation. These variables depend on the total demand for the task and the average task execution time. Since the total demand for any routine task often increases with the firm's size, larger companies typically have more viable options to consider for automation.

Generating Insights

In addition to their role in boosting productivity via task automation, LLMs are also leveraged by investment firms to generate insights and signals to inform investment decisions. For instance, hedge funds create pipelines to scan news as well as private documents related to their investments to generate actionable insights about the underlying investments.

An even more direct application of LLMs for making investment decisions is the use of near real-time unstructured data (e.g., earnings conference call transcripts) to generate systematic trading signals. These signals are then used by investment firms—particularly those using algorithmic trading—to more directly inform their trading decisions.

The use of LLM-based models for generating investment insights and signals can help firms develop a substantial competitive advantage. The benefits in such use cases depends on the firm's capability to identify right combinations of input data and processing methods that lead to outputs with high predictive power. As the identification of such opportunities is independent of the firm's size, investing resources in the development of models and pipelines for insight generation is a more promising area for smaller investment firms.

GENERATIVE AI MODELS AND PIPELINES

Many financial institutions are already highly dependent on conventional machine learning (ML) models in a wide range of operations, from credit scoring and fraud detection to portfolio optimization and algorithmic trading. While models based on conventional machine learning have proven to be invaluable for many financial institutions, powerful generative AI models present new opportunities for investment firms to enhance automation and augment human intelligence across financial applications. As institutions explore leveraging generative AI beyond the existing conventional ML models, it is useful to understand the key distinctions between developing and deploying conventional ML models versus GenAI solutions. The differences span areas such as data requirements, model evaluation approaches, workforce skillsets, and choices around model training. A better understanding of these differences can help financial firms strategically navigate the GenAI landscape and make informed decisions about how to adopt these innovative technologies effectively within their organizations.

Datasets

While the general ideas and methodology underpinning the development of LLMs share similarities with other branches of machine learning (ML), the solutions enabled by GenAI also exhibit distinct characteristics from conventional ML models. Conventional ML models typically require an extensive set of labeled data (e.g. datasets that are routinely available to online B2C platforms). Many smaller B2B businesses do not generate enough data to train ML models to effectively integrate conventional ML models into their core business activities. GenAI, on the other hand, can help boost automation and productivity in a broader array of functions across different types of organizations without necessarily requiring large datasets or even the human resources that are familiar with the art and science of ML model training and deployment. In fact, many practical GenAI solutions require the use of simple data preprocessing pipelines, carefully crafted prompts (or prompt chains), and quality control mechanisms to ensure output reliability—all of which are readily accessible to organizations with smaller data science teams and independent of the size of the datasets.

Model/Feature Evaluation

Another stark contrast between solutions that are based on conventional ML models versus those based on LLMs is the way the performance is evaluated. Conventional ML models typically have well-defined evaluation metrics (e.g., accuracy, precision, recall) that allow for a centralized quantitative assessment at the time of model development. LLM-based solutions, in part due to the open-ended nature of their interactions with users, are more challenging to evaluate. For these solutions, the evaluation and model improvement process involves collecting human feedback—ideally from those who represent different user personas—to iteratively improve the model before rolling it out to all users.

Required Skillsets

There are also meaningful differences between the skillsets required for developing successful GenAI products and those needed for conventional ML models. Training of accurate classification or regression ML models—which primarily rely on the technical capabilities of the data science teams—is often a critical step in the development of products based on conventional ML methods. The success of GenAI projects, however, often more directly depends on the capability of the product managers to develop a deeper understanding of sources and nuances of knowledge resident within the company (e.g. through collaboration with experts). This helps decide the appropriate focus areas and curate underlying information upon which the product works. GenAI product managers should have a strong sense of predicting how different personas of users interact with the product (in many cases, the input to GenAI product features is itself unstructured text queries with large variations). Finally, product managers and the technical team should have the analytical skills needed to improve model performance by compiling and analyzing scattered qualitative feedback.

In-house Training of Base Models

In-house training of large language models from scratch on a firm's private documents (and/or on a general finance-heavy corpora) can potentially be the most expensive effort in the GenAI space that a financial firm can embark on. This process embeds the firm's internal knowledge into the model's long-term memory (e.g., model trainable parameters), which, in turn, reduces the need to retrieve and include large amounts of information in the prompt during query. However, as LLMs are becoming more capable of processing larger context windows (e.g., the input sequence length) over time, a trend which is also likely to continue, financial firms may simply achieve similar results for a wide range of tasks by including more relevant information in the prompt rather than by developing (or fine-tuning) LLMs. The latter approach often requires far more expensive resources for the training and deploying (also, models often need to be trained and updated with new data over time).

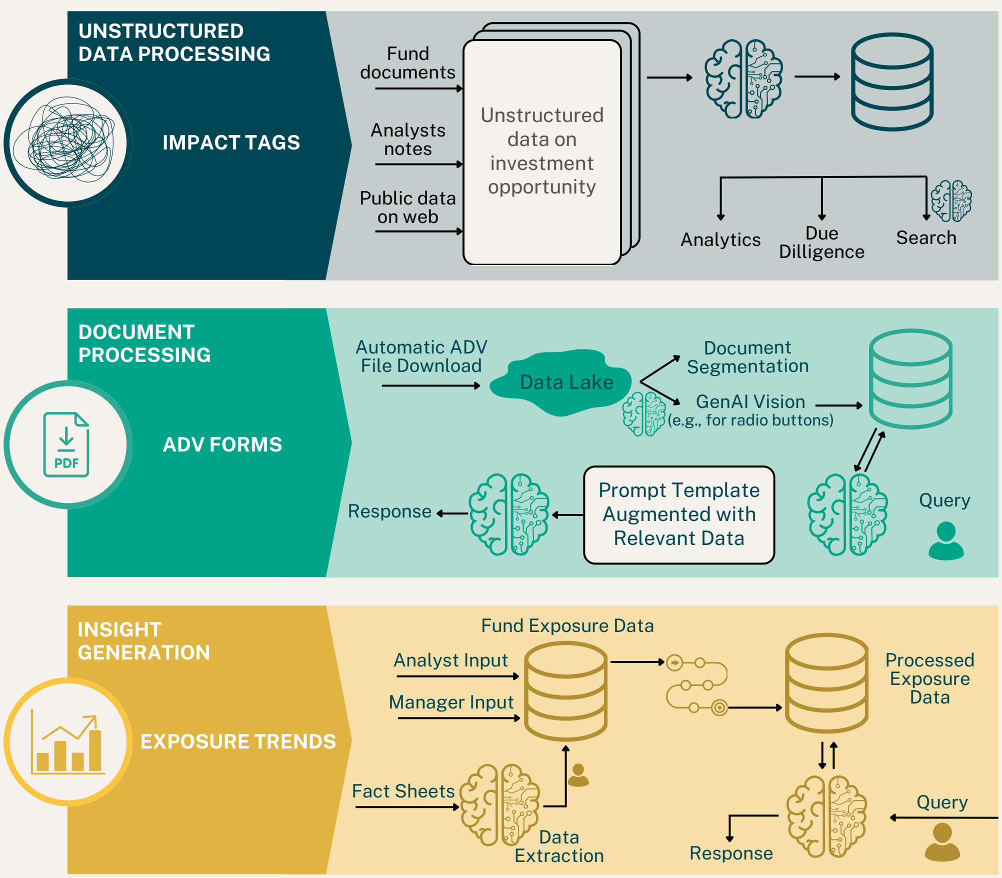
GenAI Risks

Beyond hallucination and algorithmic bias—which are broadly recognized as risks associated with the use of GenAI models—the proliferation of AI tools within organizations may also create an exaggerated notion about the depth, breadth, and objectivity of the insights [5]. Such illusions can arise from several factors, including the incompleteness of the underlying data, the AI's limited ability to consider alternative hypotheses (at least by default), the erroneous belief that AI is impartial or can represent all perspectives, as well as the tendency to conflate access to information with genuine comprehension of its meaning and implications [5]. It has also been observed that the use of GenAI in groups can adversely affect collective creativity as it leads to less diverse ideas [6].

SELECT USE CASES

In this section, we provide a brief overview of RockCreek's experiences with building customized product features and tools using LLM and integrating them in the due diligence platform. Our journey began with the development of *Asoka*, a humanoid financial assistant aimed at enhancing the analytical capabilities of our team. This initiative led to the creation of a financial assistant chatbot that incorporated both rule-based and natural language processing (NLP) models to deliver investment insights. More recently, to ensure tighter integration with existing infrastructure and utilize the latest AI models and tools, we have developed an integrated chat interface called *Sirius Chat* within our main due diligence platform. The LLM-based features implemented within the *Sirius Chat* can be broadly categorized into four groups, namely text generation, data extraction, search, and ad-hoc use cases.

Examples of LLM-based features and tools implemented on *Sirius Chat*.



Overview of select features and their associated pipelines implemented in the *Sirius Chat*.

	Use Case Examples	Main Pipeline Components	Primary Impact
Text Generation	• Compare a Fund's ADV filings from 2 periods to report important changes • Extract trends & insights from numerous periodic documents • Draft RFP responses & investment policy statements using previous documents	• Tag documents & sections by type, topic & useful tags to enhance retrieval quality, often leveraging LLMS • Use optimized, logically chained prompts to process, generate, & evaluate output, re-rank or repeat if needed • Access the firm's internal documents (in vectorized format) for information retrieval (RAG)	• Insights
Data Extraction	• Compare a Fund's ADV filings from 2 periods to report important changes • Extract trends & insights from numerous periodic documents • Draft RFP responses & investment policy statements using previous documents	• Automated file downloading & web crawling • Scanning & digitizing documents with vision tools for extracting data from tables & graphs • Optimized prompt templates & low-temperature settings for consistent LLM outputs • Locally developed LLM or affordable cloud model for large batch jobs to reduce cost • Postprocessing pipeline for data standardization, quality checks, & database/form entry	• Automation • Better datasets
Search & Rank	• Searching investment opportunities using natural language queries on structured & semi-structured datasets	• Optimized prompts extract predefined fields from unstructured queries (e.g., asset class, geographic focus, dates) • Standardize extracted fields by mapping them to standard values & expand keywords using LLM • Generate filtering logic matching data formats (e.g., Pandas DataFrames) with LLM • Use another LLM call to trim & rank candidates based on user preferences or business logic	• Insights
LLM Only & RAG	• Document summarization • Text editing • General question/answer	• Document scanning with vision tools for data extraction from tables & graphs • Access to internal documents for information retrieval	• Productivity

Design & Implementation

We started using GenAI models and tools on internal tasks with high potential impact and well-defined sources of data. For each project, besides data scientists and developers, a cross-functional team with deep knowledge of the subject matter jointly contributes to the ideation and testing of the model. We found that having subject matter experts in the loop from the inception of each project helps with design of the pipeline and prompts, which in turn increases the likelihood of broader adoption of the feature by others in the firm. As mentioned earlier, almost all the GenAI features that we have developed are implemented in the *Sirius Chat*. By incorporating other rule-based automation tools, *Sirius Chat* effectively serves as a one-stop-shop for accessing a wide range of smart tools that are developed internally to help different functions within the company.

In certain cases, after gathering adequate feedback on the performance from users, we utilize the same model (through API) across other pages of the due diligence platform. This approach helped facilitate rapid iteration and validation of the model (by avoiding the necessity for front-end development, among other factors) while maintaining ultimate consistency between different access points for the model.

One of the key lessons we learned in our efforts with GenAI is the challenge of creating a robust data preprocessing pipeline. This is particularly true in earlier stages when there was a lack of established best practices to follow and little code to recycle. Creating these pipelines typically entail vectorization of data, document processing, running large batch jobs to prepare existing data and creating a mechanism for preprocessing of incoming data in the future. Moreover, in the initial stages, significant bandwidth of the technical team was consumed in ensuring data security, enabling granular data access based on the firm's data governance rules, and setting up new software deployment practices for GenAI applications. We found that it was important to allocate enough time and resources to the creation of reusable data preprocessing code and platform groundwork, as both of these tasks play a crucial, but less visible, role in the success of GenAI solutions over the long term.

Postprocessing of Results

The effective postprocessing of the model output—particularly when the final output is compiled by putting together different pieces generated by different sections of the pipeline—plays a significant role in standardization, validation, and quality improvement of GenAI solutions. In our models, we typically run the compiled models outputs through one or more of the postprocessing steps listed below to generate the final output available to the user.

- Adding a summary of the response to the output (which itself can be generated by the LLM)
- Embedding custom graphs (by passing structured data to a general-purpose visualization function)
- Prompting the LLM to use markdown language (e.g., to generate tables when warranted by output), and provide the output in PDF (with or without a template depending on the task) and in tabular formats
- Provide references to documents and data

While current LLMs are capable of extracting insights from simple tabular data, more complex numerical calculations and analysis that might be needed within a GenAI pipeline are best handled through a hybrid approach where calculations are performed by separate functions that are called once all inputs have been compiled and cleaned. In these situations, using LLM agents can provide more flexibility as agents can make decisions about the use of auxiliary functions based on the specific context of the task being executed.

SUMMARY

Recent groundbreaking advances in generative AI has created massive hype in business settings surrounding its uses with expectations that are, in many cases, unrealistic. Successful integration of LLMs into investment operations hinges on a precise understanding of both their potential and their limitations. It is crucial for GenAI projects to stay focused on creation of tangible value for the business during the entire design and implementation process rather than pursuing nebulous and overly broad goals which routinely cause machine learning projects to fail.

At RockCreek, our work on GenAI solutions has primarily focused on creating a flexible platform to facilitate rapid prototyping and testing of new ideas. The platform, which is continually evolving, has already enabled us to create and deploy a dozen LLM-based features in our investment due diligence platform. In our experience, the modularity of pre- and post-processing pipelines plays a crucial role in streamlining the development process. Expected future advancements in the capabilities of foundation models, the availability of more sophisticated tools to leverage LLMs for various tasks, and increased exposure of technical teams to GenAI technology are three key factors that will likely accelerate and enhance the process of product development with LLMs.

For smaller financial firms with limited experience in AI, starting with an external GenAI consultant can be valuable. A consultant can evaluate potential GenAI use cases relevant to the firm, outline a roadmap for implementation, and advise on the resources and expertise needed to successfully integrate GenAI solutions. This will allow firms to explore GenAI opportunities strategically before committing significant internal resources.

In general, new firms can begin their journey with GenAI by first and foremost providing secure, private access to a fully-managed LLM service for employees through a chat interface. Adding basic capabilities that allow users to interact with uploaded documents will turn the chatbot into a sandbox for quickly iterating over different ideas and prototyping applications. Further development can initially focus on low-risk, internal-facing applications that work on one document at a time (e.g., document classification). This may then be followed by designing and incorporating more complicated applications into the platform. These typically require creating pipelines to search and access several data sources and documents and involve multiple processing steps by LLMs and their associated tools.

As the investment firms continue to explore the potential of generative AI, it is imperative to strike a balance between enthusiasm and pragmatism. While the possibilities are exciting, it is essential to ground expectations in reality and focus on the specific, demonstrable value that GenAI can bring to the operations. By embracing a more nuanced and precise approach, the industry can harness the true power of these remarkable technological advancements while avoiding the pitfalls of unrealistic expectations and overselling.

REFERENCES

1. The State of AI in Early 2024: Gen AI Adoption Spikes and Starts to Generate Value, McKinsey, May 2024.

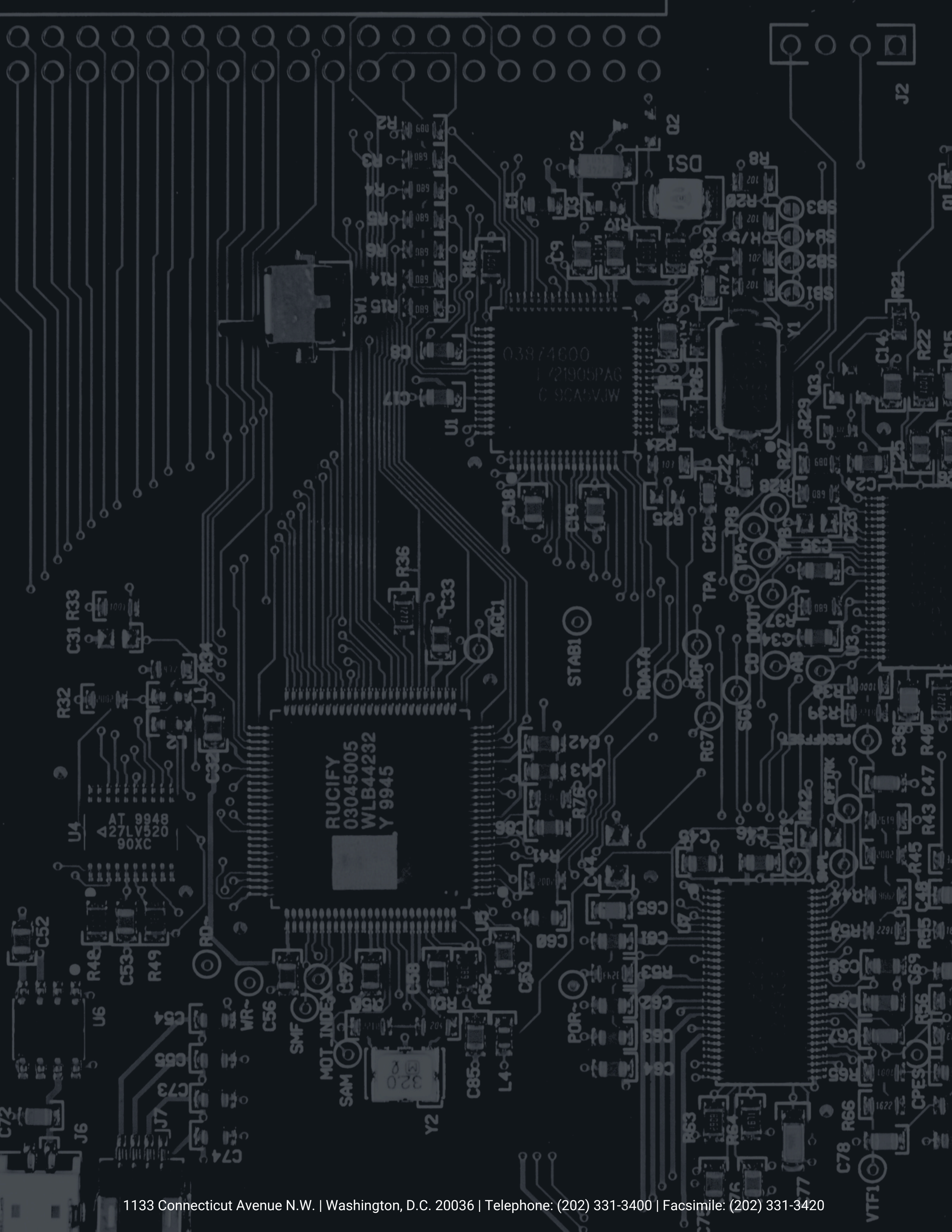
2. Generative AI: The Insights You Need from Harvard Business Review, January 2024.

3. A Generative AI Reset: Rewiring to Turn Potential into Value in 2024. McKinsey Quarterly, May 2024.

4. The AI Index 2024 Annual Report, AI Index Steering Committee, Stanford University, Stanford, CA, April 2024.

5. L. Messeri, M.J. Crockett, Artificial intelligence and illusions of understanding in scientific research, Nature 627, 2024.

6. How People Can Create—and Destroy—Value with Generative AI, BCG Henderson Institute, October 2023.



1133 Connecticut Avenue N.W. | Washington, D.C. 20036 | Telephone: (202) 331-3400 | Facsimile: (202) 331-3420