



Automating Counterparty Credit Risk A Case Study on GenAI in Asset Management

By Peter Griffiths, Murali Menon

Building a Universal Client-Centric Solution

Automating Counterparty Credit Risk: A Case Study on GenAI in Asset Management, is a proof of concept (PoC) involving a European asset management firm with over €30 billion in assets under management (AUM). The firm needed to automate and enhance its counterparty credit assessment process. Recognised as advocates and innovators in AI/ML, and having collaborated with early adopters in financial services, Calimere Point were chosen as key partners to conduct a comprehensive PoC. Our objective was to evaluate the potential of leveraging generative AI solutions. In this case study, we examine the implications of the generative AI (Gen AI) revolution, address existing challenges in counterparty credit assessment, demonstrate real-world application and benefits of the PoC, and validate the impact of the findings on asset and wealth management.

Exploring Generative AI's Potential: Generative AI holds immense potential for asset and wealth management. Understand the cultural and procedural shifts needed to unlock its capabilities.

[Connect on LinkedIn](#)

Contents

1. Introduction
2. Overview and Context
3. Addressing Key Challenges
4. Insights and Discoveries
5. Experimentation and Approach
6. Overall Conclusion and Analysis
7. Opportunities with Generative AI

1. Introduction

Navigating the GenAI Revolution: Implications for Asset and Wealth Management

The generative AI (GenAI) revolution is well underway, promising exceptional performance in information search, retrieval, and synthesis tasks on unstructured content, as well as content generation capabilities across text, images, and code. This capability to process vast amounts of information, synthesise contextually, and auto-generate human-friendly responses is transforming the way we live and work.

As a transformative ally, GenAI is emerging as a significant disruptor for asset and wealth managers. Innovative firms are already deploying GenAI and reaping tangible benefits across a broad array of use cases by investing in the necessary infrastructure and talent to achieve significant efficiencies.

What Does Gen AI Mean Specifically for Asset and Wealth Management?

Organisations who are still considering GenAI adoption are being left behind—the time to implement is now. However, it is crucial to recognise that despite the hype surrounding GenAI, there are still unknowns, barriers to adoption, and limitations with the toolsets.

Unknowns

Legal Implications: The legality of outputs generated by foundational models remains uncertain, raising questions about intellectual property and compliance.

Accuracy and Bias: GenAI can produce inaccurate results, hallucinations, and may contain inherent biases, with a lack of explainability for its decisions.

Data Security: Ensuring the safety and privacy of client data is a major concern, especially with increased regulatory scrutiny.

Internal Compliance: The use of GenAI approaches present a challenge to the internal compliance functions at firms - getting compliance teams comfortable with the application of GenAI requires significant organisational investment to upskill all functions.

Regulatory Compliance: The use of GenAI must not violate regulatory or fiduciary rules, necessitating thorough vetting and monitoring.

Barriers to Adoption

Talent Shortage: There is a significant lack of skilled professionals capable of developing and implementing GenAI solutions.

Data Quality: Lack of data foundations, as poor data quality remains a challenge, hindering the effective deployment of GenAI in asset and wealth management.

Cybersecurity Threats: The models and training data are vulnerable to cybersecurity threats, necessitating robust protective measures.

Investment Costs: The high costs associated with implementing GenAI, including infrastructure and ongoing maintenance, can be prohibitive.

Toolset Limitations

Contextual Understanding: Gen AI tools may struggle with understanding complex contextual nuances, which can lead to inaccurate or irrelevant outputs.

Scalability Issues: Scaling GenAI solutions across large organisations can be difficult, requiring significant computational resources and infrastructure.

Integration Challenges: Integrating GenAI with existing systems and workflows can be complex and time-consuming, often requiring custom solutions.

User Training: Effective use of GenAI tools necessitates comprehensive training for end-users, which can be resource intensive.

Ethical Concerns: Addressing ethical concerns, such as the potential for misuse or unintended consequences of AI-generated content, is critical for responsible implementation.

Introduction Summary

As AI technologies evolve and become more widespread, new challenges and risks will undoubtedly arise. Our role is to help organisations establish responsible AI programs that ensure ethical usage and safeguard against potential disruptions.

We understand the immense potential of generative AI and the necessary cultural and procedural shifts required to unlock its full capabilities. Furthermore, we prioritise educating our clients about the inherent risks of AI and methods to mitigate these risks, thus preventing harm to customers, businesses, and society.

2. Overview and Context

About the PoC

The case study cuts through the hype to present facts and key limitations of generative AI tools, along with alternative approaches. It aims to demonstrate practical applications for asset and wealth managers. By working with early adopters in this space, we have helped them evaluate the viability, usability, and feasibility of generative AI features and systems.

Data Analytics Team

The GenAI in Asset Management PoC was delivered by Calimere Point, a London-based data analytics consultancy. Calimere Point is a leader in developing and delivering powerful, bespoke data analytics and visualisation solutions. We have successfully implemented solutions incorporating both predictive and generative AI models. For 15 years, investment banks, asset managers, and financial organisations have relied on our team to gain deeper insights, enhance decision-making, and enable rapid development and deployment within complex data landscapes. To date, Calimere Point has generated over £150 million in proven revenue uplift and cost reduction benefits for our clients. One of our firm's strengths is our consultants, who combine data science expertise with direct market experience across Investment Banking (trading/structuring/risk/finance) /Asset Management and other industry verticals. We leverage a combination of senior-level business experience and cutting-edge data analytics capabilities to drive value within our clients' organisations.

Scope

A prominent European asset management firm, with over €30 billion under management (AUM), engaged Calimere Point to conduct an in-depth proof of concept exercise. The goal was to assess the feasibility of leveraging generative AI solutions to automate the counterparty credit assessment process. While many wealth and asset management firms have already integrated AI into their core operations, the maturity of GenAI use cases is still evolving. Therefore, it was crucial for the firm to select the right partner and have confidence in the project's viability and success criteria.

Background

The asset management firm completed a comprehensive annual credit assessment of its 150 trading counterparties, a crucial step in its Investment Risk process. Among these, 100 were unlisted entities.

For the listed counterparties, they were able to leverage regulatory filings to drive the credit assessment, however for the non-listed firms, they had to depend solely on the firms' annual reports.

To tackle this challenge, they mobilised internal resources to meticulously review and extract relevant financial information from over 100 annual reports. Despite their efforts, the manual review process posed significant challenges and inefficiencies.

Conducting detailed annual credit assessments of 150 trading counterparties is a multifaceted challenge that requires substantial resources, sophisticated analytical tools, robust data management practices, and strict adherence to regulatory standards.

The Client Inquiry: Addressing Operational Inefficiencies and Document Diversity

In response to challenges in their Investment Risk Process, the asset management firm needed solutions to streamline operations and enhance efficiency.

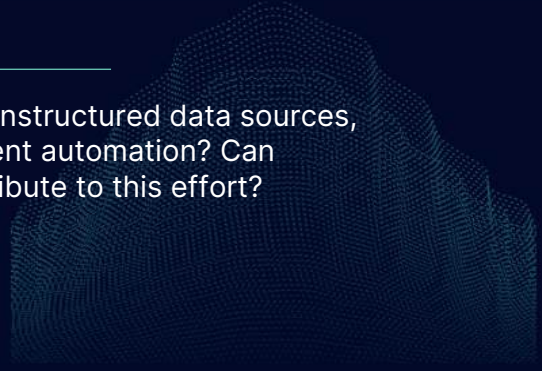
The following list outlines the primary needs identified by the client as they navigate the complexities of data validation and processing in a diverse document landscape:

How can we optimise time and cost resources, especially with skilled personnel engaged in repetitive manual tasks?

What strategies exist to minimise the time spent on validating data, considering potential human errors in the current process?

Is automation feasible for our manual-intensive processes, given the diverse structures and formats of documents? Can generative AI models aid in overcoming this obstacle?

Considering the complexity of unstructured data sources, how can we effectively implement automation? Can generative AI significantly contribute to this effort?



3. Addressing Key Challenges

1. Manual Review Challenges

Time-Consuming and Labour-Intensive

The manual review process was time-consuming and labour-intensive, resulting in significant cost issues and high resource demands. Highly skilled and expensive risk staff were required to perform repetitive tasks manually.

Data Extraction

Analysts had to sift through extensive annual reports, often spanning hundreds of pages, to extract pertinent financial information. This approach not only strained resources but also had a high potential for human error.

Validation

The manual process required a rigorous set of checks and controls which consumed further resources. Sample testing approaches didn't guarantee the accuracy of the complete population.

2. Existing Counterparty Credit Process Challenges

Sourcing

Searching for reports is tedious and manually intensive:

- For US companies: EDGAR database
- For UK companies: Companies House database
- For other geographies: ad-hoc search on corporate websites or local regulator pages

Data Extraction

- Manually open each report and look for total assets / total equity values in balance sheet or statement of annual position pages
- Keep track of reporting currency and possible multipliers (e.g. \$'000s)
- Copy-and-paste values into Excel

Variability

- Counterparties may change names over time, making it hard to keep track.
- For EDGAR, need to look up company identifiers (CIK), but the look-up is not robust

3. Risk Management Implications

Delayed Analysis

Delays in accessing and analysing financial information could hinder the firm's ability to identify and mitigate potential credit risks in a timely manner, leaving it vulnerable to market fluctuations and financial downturns.

4. Annual Reports Data Structure – Managing Inconsistency

Inconsistent Formats

Reports are in PDF, not machine-readable (e.g., XBRL). PDFs often contain embedded images, which are not text-searchable.

Equity and Assets Measures

- Equity and assets measures usually appear in tables
- Sometimes, values appear without headings
- Misaligned values and multiple columns for different years
- Total equity not always explicit (e.g., Shareholders' or members' equity, Members' interest, Stakeholders' funds)
- Total assets not always explicit (e.g., Fixed assets + current assets, Fixed assets = non-current assets).

4. Insights and Discoveries

Potential for Automation: A Critical Finding

Assessment Effort

The analysis revealed that the entire annual Counterparty Credit Assessment process consumed approximately 30 days of effort each year. This extensive time investment underscores the pressing need for efficiency improvements.

Complexity of Unstructured Data

Prime target for automation, the process faces significant hurdles due to the complexity of unstructured data sources. Manual extraction and validation processes struggle to cope with the volume and variability inherent in these sources.

Exploring Gen AI Solutions

Given the challenges posed by the complexity of the data, an intriguing question arises: Could generative AI offer a viable solution? How might it streamline data extraction and validation processes while maintaining accuracy and reliability?

Strategic Considerations

As we deep dive into the potential of Gen AI, it's essential to consider the strategic implications. How might implementing such technology impact resource allocation, risk management practices, and overall operational efficiency?



Basic Questions and Challenges

Can a large language model help with text extraction?

Exploring the potential of leveraging advanced language models to streamline text extraction processes represents a fundamental query in our pursuit of efficiency.

Given a PDF, can a query such as 'extract total assets and total equity from the attached company annual report' find the metrics?

The ability to precisely extract critical financial metrics from PDF reports through tailored queries is a pivotal consideration in enhancing automation and efficiency.

Is a Programmatic interface available?

The availability of a programmatic interface capable of seamlessly integrating with existing systems emerges as a pivotal aspect in the quest for streamlined and automated processes.

5. Experimentation and Approach

Generative AI is transforming the financial services industry, particularly in asset and wealth management. For firms to harness its full potential, a dual focus is essential: understanding the technology and taking decisive action while managing associated risks judiciously.

Value creation from GenAI will not only stem from advanced technology but also from cultivating a data-centric culture that invests in foundational capabilities and develops robust risk management frameworks. Successful initiatives will emerge from a blend of industry domain expertise and a culture of innovation, envisioning new ways of doing business through the convergence of GenAI with traditional data analytics and hybrid approaches.

Summary of Approaches

We explored three different approaches to automating counterparty credit assessment:

1. **Pure Generative AI:** Focused on leveraging advanced AI for interactive and natural language processing tasks.
2. **Traditional Data Analytics:** Emphasised accuracy, repeatability, and robustness through established data processing techniques.
3. **Hybrid Approach with AI Assistants:** Combined the strengths of generative AI and traditional data analytics, ensuring control, accuracy, interactivity, and scalability.

Key Insights from Initial Findings

1. **Platform Capabilities:**
 - Claude by Anthropic and Gemini from Google: Facilitated PDF uploads.
 - ChatGPT by OpenAI: Did not support PDF uploads at the time of the PoC
2. **Process Limitations:**
 - Constraints such as file size, maximum number of files per session, and token number restrictions:
 - ChatGPT: Approximately 4,096 tokens.
 - Claude: Up to 100,000 tokens.
3. **Query Management:**
 - Each follow-up query requires recomputation of previous queries, rapidly exhausting the token limit.

Recommendations for Optimising Generative AI Use

1. **Consolidate Instructions:** Combine multiple instructions into a single query to optimise token usage. This allows the language model to distil pertinent aspects of combined queries.
2. **Caution:** This approach may not always yield desired outcomes as some instructions might be overlooked.
3. **Address Non-Deterministic Responses:** Be aware that response outcomes for identical input data and queries can be non-deterministic, requiring verification for consistency.

Approach 1: Pure Generative AI

Introduction to Claude by Anthropic

Anthropic, an AI startup based in San Francisco, California, describes itself as an AI "safety and research" business. The company focuses on creating "reliable, interpretable, and steerable AI systems." Google invested nearly \$2.4 billion in Anthropic, and Amazon pledged up to \$4 billion. Anthropic primarily focuses on building large language models (LLM). In 2024, the company's most advanced language model is Claude 2.1, a sophisticated AI assistant similar to ChatGPT.

Claude Execution Approach

File Upload Capabilities

Claude allows the upload of up to five files at a time, each less than 10MB. While manual upload is available, programmatic upload options are not currently provided.

Evaluation Query

Extract into a table the company name, registration number, as of date, fixed assets, current assets, total equity, and the detected currency in ISO format from the attached financial statement PDFs. Take the registration number from the document file name. Create a new field called total assets as the sum of fixed assets and current assets. Note that in some financial statements, fixed assets appear as non-current assets, and total equity can appear as members' equity, stakeholders' equity, members' interests, or similar. Remove currency symbols from the fixed assets, current assets, total assets, and total equity values. Adjust the output numbers appropriately if the financial statements state that the numbers are in thousands or millions.

Sample Results

When it works, it's excellent

- Name, registration number, as of date, currency correct
- Assets and equity also correct 4 out of 5 times
- Worked with non-English languages as well

Danger! When wrong, it's often subtle

- MUFG Bank values ignore 'millions of yen' note
- Seaport, Park Walk total equity is incorrect; ignored the instruction that "total equity can appear as members' equity or stakeholders' equity or members' interests or similar."

Results Consistency Issues

The most serious problem is the lack of repeatability. Running the same query with the same inputs does not guarantee the same outputs. The table below shows the percentage of times, out of five runs, that the same value was found. Each run requires verification, which, while potentially faster than manually opening PDFs and copy-pasting values, is also prone to human error. Thus, the benefits of using LLMs in this context are questionable.

General Issues

- Requires iterative instructions to correct equity and multiplier, each of which further consumes the token limit.
- When results for five uploaded files are correct, upload the next five and ask to "update table with attached set of files."
- Results need to be checked for each upload.
- Likely to run out of token limits long before processing 100 companies.
- Even a professional subscription (500K tokens) is probably insufficient for processing all companies.
- Files larger than 10MB cannot be processed; OCR'd files are generally much larger than 10MB.
- A pure LLM method is not scalable.

Conclusion

In summary, Claude demonstrates potential as an AI assistant with robust capabilities for processing and extracting data from financial statements. When functioning correctly, it excels in accurately interpreting and extracting relevant information, even from non-English documents. However, its performance is inconsistent and often requires manual verification due to issues with repeatability and subtle errors. The current limitations in file size, token limits, and lack of programmatic upload options hinder its scalability for large-scale applications. While promising, further improvements are necessary to make Claude a more reliable and scalable solution for extensive data processing tasks.



Approach 2: Traditional Data Analytics

Introduction

To provide a comprehensive evaluation, we not only tested generative AI models but also leveraged traditional data analytics approaches to address the same challenge. This allowed us to compare the effectiveness, accuracy, and scalability of both methods directly.

Automated Download Process

For the traditional data analytics approach, we automated the process of downloading relevant financial documents from public databases using Python packages:

EDGAR (US Companies) API:

Lookup CIK Given Company Name: Utilise the EDGAR API to retrieve the Central Index Key (CIK) using the company name, essential for querying specific filings.

Find Documents by Filing Type: Search the EDGAR database for documents of specific filing types associated with the CIK.

Annual Report Filing Type: 'X-17A-5': Target annual reports by filtering for the 'X-17A-5' filing type.

Download PDF of Document: Automate the download of the annual report in PDF format for further analysis.

Companies House (UK Companies) API:

Search for Company Data: Leverage the Companies House API to search for data on UK companies.

Obtain Latest Filings for 'Active' Status: Retrieve the most recent filings for companies that are currently active.

Download PDF of Document: Automate the download of the relevant documents in PDF format. Given that many financial statements are embedded in images within PDFs, we applied OCR (Optical Character Recognition) to extract and overlay text onto these documents, making them searchable.

Programmatic Pre-filter

Extract Metrics via Text-Based Programmatic Search Where Possible First:

Python Package to Read PDF Files: Use Python libraries such as PyPDF2 to programmatically read and extract text from PDF files.

Regular Expressions to Find Pages with Required Patterns: Deploy regular expressions to locate and extract pages containing specific patterns related to the required financial metrics.

Will Fail if Required Value Has No Corresponding Heading: This method depends on identifiable headings and will fail if such headings are missing.

Will Fail if No General Pattern Found Across a Set of Reports: Consistency in report formatting is necessary for success; the approach fails if no general pattern is detected.

Non-English Language Reports are Difficult to Handle: Handling reports in languages other than English presents significant challenges due to language-specific formatting and terminology.

PDF Optimization

Optimize PDF File Size:

Compression Methods: Implement various compression techniques to reduce the size of PDF files. While this can help manage large files, it may lead to data quality loss, presenting a trade-off between file size and data integrity.

Conclusion

The use of traditional data analytics approaches enabled us to construct a robust and repeatable automated process quickly and efficiently. This process was highly accurate and provided consistent results with each execution, a level of reliability that the generative AI engines could not achieve.

Traditional methods ensured that each step of the data extraction and processing workflow was controlled and verifiable, highlighting the strengths of a more deterministic approach to handling financial document analysis. This approach proved to be a dependable alternative, offering the asset management firm confidence in the feasibility and reliability of automating their counterparty credit assessment process.

Approach 3: Bespoke Hybrid Model with AI Assistants

Introduction

In November 2023, OpenAI introduced a new feature in its ChatGPT generative AI solution called the ChatGPT assistant. This innovation allows for the hybridisation of pure generative AI models with traditional data analytics solutions.

This new hybrid model was explored as part of our PoC to combine the interactivity and accessibility of generative AI with the accuracy and consistency of traditional data analytics.

Overview of AI Assistant Functionality

Our earlier exercises demonstrated that generative AI models alone struggled to accurately and consistently extract company information. On the other hand, traditional data analytics approaches provided more accuracy and consistency but lacked the interactive element offered by generative AI interfaces.

How AI Assistants Enhance Data Processing

AI assistants provide a hybrid approach that combines the benefits of generative AI models and traditional data analytics. The AI assistant integrates traditional data analytics logic, which is executed in the background, while the generative AI interface facilitates user interaction.

Benefits of the Hybrid Approach

- **Provides Guardrails:** Enhances generative AI models' accuracy and consistency through controlled, robust data analytics.
- **Interactive User Interface:** Maintains the high interactivity level offered by generative AI interfaces.
- **Versatile Logic Integration:** Allows embedding of simple or complex data analytics processes, providing flexibility in functionality.

Example (simplified) – LLM running SOLO - Performance without AI Assistants

Simple Query:

Question: Who is the President of the United States of America?

Response: ChatGPT accurately answers that Joe Biden is the President of the USA.

Complex Query:

Question: Extract the financial information for Mizuho Securities LLC.

Response: Left to its own devices (without assistance), ChatGPT struggled to extract accurate and consistent data for financial metrics for Mizuho Securities LLC, despite being spoon-fed the data via a specific text extract. After manually uploading specific text from the PDF of Mizuho Securities LLC's financial statements and refining the query, ChatGPT could occasionally extract the relevant metrics, but responses varied and were volatile.



Example (simplified) – Performance with AI Assistants

Incorporating an AI assistant changes this scenario:

Enhanced Query Handling:

Question: Extract the financial information for Mizuho Securities LLC

Response:

- The assistant allows us to put a set of pre-scripted code, designed to carry out specific tasks, behind the ChatGPT user interface
- The user interacts with the ChatGPT interface in the same way as before, but using specific prompts the interface invokes the assistant rather than relying on the LLM to provide the response for that specific query.
- The answers are returned by the ChatGPT interface for both the LLM response and the Assistant response.

Advantages of Using AI Assistants

Control and Accuracy: Ensures accurate and repeatable responses through predefined functionality.

User Interactivity: Maintains high levels of interactivity offered by generative AI models.

Scope for Integration: Allows for unlimited integration of AI assistants, enabling complex guardrail infrastructure without compromising interactivity.

Ease of Deployment: AI assistants are easily built and deployed.

Auditability: The code for AI assistants can be exposed and audited to validate accuracy.

Complex Logic Construction: Enables embedding of generative AI into advanced processes for sophisticated functionality.

6. Overall Conclusion and Analysis

Our PoC exercise evaluated three approaches to automating counterparty credit assessment: using generative AI models, traditional data analytics, and a hybrid model with AI assistants.

1. Pure Generative AI:

Strengths: High interactivity and potential for natural language processing.

Weaknesses: Inconsistent results, lack of repeatability, and scalability issues.

2. Traditional Data Analytics:

Strengths: High accuracy, repeatability, and robustness.

Weaknesses: Lack of interactivity and the need for extensive pre-processing and coding efforts.

3. Hybrid Model with AI Assistants:

Strengths: Combines the benefits of both generative AI and traditional data analytics. Ensures control, accuracy, interactivity, and scalability. Facilitates complex logic integration and offers auditability.

Weaknesses: Requires initial setup and integration efforts to develop and deploy AI assistants.

Conclusion

Calimere Point brings together all the components of a solution. We built a bespoke Hybrid model with AI assistants that emerged as the most effective solution, offering the robustness of traditional data analytics with the interactive, user-friendly nature of generative AI. This method provides a scalable, accurate, and consistent way to automate counterparty credit assessments, aligning well with the firm's requirements and demonstrating a viable path forward for integrating advanced AI into asset management processes.

To recap what it took:

2 data scientists

1 month

10x cost reduction

7. Opportunities with Generative AI

Given how rapidly Generative AI has taken off, the initial step is to begin with foundational education, ensuring that everyone—including the Board, C-suite, and staff—understands its importance. This is crucial for creating a level playing field across the organisation. Companies that invest time early will be better positioned to navigate GenAI most effectively.

All asset managers will need a robust GenAI strategy to maximise benefits and minimise risks. GenAI can be a transformative breakthrough for asset managers at a critical time for the industry. The natural tendency might be to adopt a "wait and see" approach, as with most technology trends. However, we believe that waiting is not an option due to the rapid growth of GenAI. Asset managers urgently need a forward-looking GenAI strategy to move forward with confidence, mitigate risks, and reap the benefits of this powerful new technology.

Calimere Point, a strong advocate of AI's capabilities, is uniquely positioned to assist organisations in leveraging generative AI. With our extensive background in machine learning and natural language processing, we support organisations through every stage of AI integration. Our Hybrid model is live and currently in use. Learn how Calimere Point can help make your generative AI implementation successful and explore how we can assist you in adopting AI.

Contact

Peter Griffiths | CEO

Direct: +44 203 290 7817

Mobile: +44 7787 540764

peter.griffiths@calimerepoint.com

Murali Menon | Head of Data Science

Direct: +44 203 290 7817

Mobile: +44 7952 358194

murali.menon@calimerepoint.com



Calimere Point is a data analytics consultancy based in the United Kingdom. We provide our clients with value by developing robust data analytics capabilities that enable deeper insight and better decisions.

<https://calimerepoint.com>