JS|HELD

PERSPECTIVES

Leveraging Generative AI in Complex Investigations

Our perspectives feature the viewpoints of our subject matter experts on current topics and emerging trends.

INTRODUCTION

Complex investigations are on the rise with investigators being asked to do more with less, therefore needing to gain efficiency to handle workloads. A complex investigation can be triggered by a litigation event, government agency mandate, internal compliance activity, or corporate transaction. Investigations can encompass any number of subjects and many are done privately, but in terms of fraud probes, the Association of Certified Fraud Examiners (ACFE) estimates $5 trillion is lost globally each year to fraudulent schemes alone.

More than 75% of corporate law departments are seeing an uptick in matter volume in 2024. Yet, with more than half reporting flat or decreasing budgets, they are looking for technology to do more with less. The need or better technology is further exacerbated by the growing volume of data in the corporate landscape. Some 41% of organizations now manage at least 500 petabytes of data (a petabyte is a million gigabytes).

Additionally, users are creating roughly 25% more data each year than the previous year. Investigations can take a toll on budgets when it's been estimated that discovery costs can be as high as $12,000 to $18,000 for each GB of data reviewed. It is becoming clear that the only way for legal teams to combat the growing amount of data is by leveraging technology.

Whether examining fraud, bribery, cybercrime, cartel activity, intellectual property theft, regulatory violations, or other areas, investigators need better ways to sift through vast quantities of data to find the information at the heart of the matter. This article examines a more efficient approach to complex investigations, through the use of Generative AI (GenAI). The information in this article is intended to assist general counsel, outside law firms, forensic accountants, and other internal investigation teams, in determining the effectiveness of using GenAI in fraud and other complex investigations.

BENEFITS OF GENAI IN FRAUD DETECTION & CYBERSECURITY AI SOLUTIONS

Using GenAI Q&A chatbot interfaces, a small team of investigators can search a massive document set quickly to get answers or, at the very least, important takeaways to the most crucial questions. Rather than manually searching vast document sets, deploying keyword searches which often return too many false positives, or implementing techniques that require a lot of pre-work, investigators simply can ask questions via a Google-like search bar. Because there are no complex processes to explain or syntax learn, the investigatory team gets right to work uncovering the facts.

The system generates answers to questions, but the responses should always be taken with a grain of salt as the intelligence, while informative, is still "artificial." Although the responses are insightful, the GenAI Q&A chatbot tool refers the investigators directly to the underlying reference documents used to generate the answers, which are ranked in order of relevance.

Consider this recent project. A small team of investigators spent six weeks searching a database of 1.5 million documents and had not yet found the answers they were seeking. The client was hesitant to add additional team members to the document review task. However, interviews with key individuals were looming, and the team needed to find relevant documents for questioning. The investigation team utilized the GenAI Q&A chatbot to quickly identify areas of interest and the most relevant documents for each topic and interviewee. Within 15 minutes of using the chatbot, it uncovered 50 highly relevant documents to the investigation, with 80% of them not previously found in keyword searches.

Keyword lists are typically generated at the outset of the matter when many facts are not yet known and tend to stay static, thus not moving at the same speed as the investigation. This issue may cause an investigations team to fail to identify crucial information for the matter. Introducing GenAI-powered chatbot interfaces may not only uncover key documents earlier, but also allows investigators to pivot quickly as the matter evolves (i.e. due to a revelation from an interviewee or new questions from a regulator), and answers to a dynamic set of new questions are needed.

REDUCING TIME AND COSTS WITH GENERATIVE AI

Being able to find the most relevant documents and information at the commencement of an electronic discovery review has major implications for downstream time and cost. Knowing key facts at the outset of a review rather than after it is completed can save what might be months of document review. Also, by eliminating the issues the investigation "might be about" and focusing in on the crucial questions, such as who, what, when, why, and how, that affect the investigation's outcome allows investigators to spend time zeroing in on a richer and narrower set of documents, and consequently provide higher quality results during the review process.

Having this technology at an investigation team's disposal allows them to begin the matter with a smaller review group, reducing hourly document review costs by what could be hundreds of thousands of dollars on large matters. For matters with multiple workstreams, separate teams of reviewers can be used to focus on each workstream in parallel. With the flexibility GenAI provides in asking for any type of information in seconds, smaller dedicated teams can handle multiple issues and workstreams, further lowering reviewer headcount and cost.

Another cost driver affecting investigators is the collection, processing, and hosting of large data volumes which are only increasing as employees create and maintain more and more data. Similar to the keyword list problem of being overly broad at the outset so as not to miss anything, many times too much data is loaded prior to review. This

can result in a costly ramp-up period. With better tools to learn about a matter early on, investigators can take an approach of collecting and preserving the same wide swath of data sources as before, but only initially loading data from high priority individuals or sources until there is a sense of where the investigation will go.

In another example involving an alleged fraud scheme within an accounting department, our team collected and preserved data from over 20 potentially relevant individuals and sources. But we only loaded five custodians' data into the eDiscovery review platform powered by GenAI searching. After examining the loaded data via the chat interface, we got a better sense of the narrative and the key players. It was determined we only needed to load two more of the preserved data sources, leaving more than ten data sources "on the cutting room floor" as that information was not required to complete the investigation. This limitation spared the need to process and host multiple terabytes of data, thus saving hundreds of thousands of dollars in data-related fees alone. Performing a document review in conjunction with an investigation can be a daunting task. Therefore, reducing the number of relevant documents at the outset allows the team to identify "hot" documents more efficiently.

ENHANCING INTERVIEW PREPARATION WITH GENERATIVE AI INSIGHTS

No matter how much we can learn from the data, a crucial part of any investigation remains interviewing key individuals involved. A successful interview hinges on maximizing the time available with the interviewee by prioritizing the questions and topics and being well versed with the facts to avoid being caught off guard with surprises. In many instances, investigators only get one opportunity to question an individual, magnifying the importance of adequate review and preparation of topics and questions.

As noted, GenAI chatbot tools enable investigators to learn key information quickly and help them better understand topic areas for interviews. They can also

directly ask the technology to "provide us with a list of ten interview questions" for each person of interest. |Because the list of questions is generated quickly, investigators can adjust the prompt request to improve the quality of the questions. For example, investigators can draft questions to home in on certain topics – the more specific the better – and focus on communications with other specific individuals or stick to a specific date range. The prompts can also be designed to create lists of questions from the perspective of various roles. Asking what 10 questions an "insurance adjustor" might ask would differ from questions a "forensic accountant" or a "law enforcement officer" might have. By prompting |the GenAI chatbot tool for questions from different viewpoints and varying the phrasing, investigators can build a robust interview strategy in less than an hour.

Once the interview questions are assembled, with a combination of output from the GenAI and human refinement, investigators can then send the questions again as input to the GenAI chatbot, asking "how would the interviewee answer these 10 questions?" The answers returned may help refine the questions even further and are also hugely helpful in preparing for the interview itself. Think of it as having a free 'practice interview' with a person's data footprint before speaking with the person directly. This tactic is another way to minimize surprises and help get the most out of the interview.

STREAMLINING EXPERT REPORTS WITH GENERATIVE AI ASSISTANCE

With all of the research, investigation, and interviewing complete, investigations teams are often requested to draft an expert report summarizing the investigation process and procedures, analysis, and findings. Depending on the nature of the case and circumstances, the client or users may request interim reports or reports outlining different workstreams. Previously, eDiscovery platforms have aided investigators in identifying key facts and documents, but not in communicating that information in the form of a report.

By its very nature, GenAI is generative and can now help in this important step. To be clear, the experts will always have ownership of their reports and at the present time will still have to thoroughly review and edit the report. After all, the consequences of submitting generated AI-written content without human checking are well documented.

However, GenAI provides valuable assistance to experts by utilizing a combination of coding, investigatory notes, and document level AI summaries to build a sort of "summary of summaries," building a draft report which conveys the relevant document content. To buttress the report narrative, the actual custodian documents can be attached as exhibits, further substantiating the findings. A key tenet in using AI is to always reference the actual content from which the findings were derived – trust but verify. Additionally, GenAI can create event timelines to summarize activity across all documents and be included as another key exhibit, which is often required in expert reports. The technology generating the content is improving rapidly, and the better it gets, the faster investigators can produce an accurate and complete report.

INTEGRATING GENERATIVE AI AS A COMPLEMENTARY TOOL IN FRAUD INVESTIGATIONS

With all of the benefits that can be achieved quickly with minimal costs (or better yet, cost savings), it's not surprising that GenAI is being used to power investigations. However, that doesn't mean other data-driven techniques used in the past have been rendered obsolete. Technology Assisted Review (TAR), predictive coding, concept clustering, sampling, and other tools, are still being used. In fact, it can be said that GenAI works well as a complementary piece in an analytical toolkit. For example, the few dozen most relevant documents returned in response to a question asked of a GenAI chatbot can be used as seed documents in a predictive coding workflow or can be examined in a visual clustering interface to find documents related to the same concepts.

CONCLUSION

The use of GenAI can streamline fraud and other complex investigations and deliver reductions in costs and time spent by narrowing down the most relevant information and documents, pinpointing the individuals to interview, enhancing interview topics and questions, and helping to generate more precise summary reports. Above all, successful investigations require skilled investigators who can use this emerging technology properly. The role of technology is uncovering the facts. In our experience, GenAI helps find more of the facts, more efficiently.

ACKNOWLEDGMENTS

We would like to thank our colleagues, Mike Gaudet and Natalie Lewis, for providing insight and expertise that greatly assisted this research.

Mike Gaudet is a Managing Director providing Digital Investigations & Discovery services within J.S. Held's Global Investigations Practice. He has more than 20 years of experience providing solutions for corporations, legal teams, and government agencies related to data discovery and governance challenges. He is an expert eDiscovery practitioner and technologist, with a master's degree in computer science & engineering. He has proficiency in leveraging the right tools to quickly gain insight from data, and to efficiently achieve project goals on time and under budget.

Mike can be reached at mike.gaudet@jsheld.com or +1 281 415 5742.

Natalie Lewis is a Senior Vice President in J.S. Held's Economic Damages & Valuations Practice. As a Certified Public Accountant and Certified Fraud Examiner, who is also certified in financial forensics, Natalie specializes in forensic accounting and the analysis of economic damages. She provides consulting and expert services for both plaintiff and defense law firms throughout the country. Natalie has conducted fraud investigations for companies of all sizes as well as government entities. Her investigations have spanned the globe and include various internal investigations, embezzlement, Ponzi schemes, commercial crime insurance claims, Foreign Corrupt Practices Act (FCPA), and asset misappropriation.

Natalie can be reached at natalie.lewis@jsheld.com or +1 470 852 4601.