CAN GENERATIVE AI GIVE US PRESCRIPTIVE ANALYTICS?

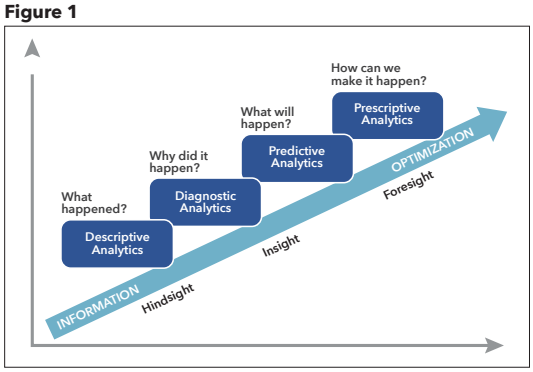
One of four key types of analytics has long been considered a "pie-in-the-sky" concept for fraud investigators. Rather than describing or diagnosing something that's happened or predicting what could happen, prescriptive analytics can tell us what we should do about it. With the growth of generative AI and large language models, it may just be in reach.

COLUMNIST
VINCENT M. WALDEN, CFE, CPA
CEO, KONA AI

When I was a partner at one of the Big Four accounting firms, clients asked me to provide recommendations and guidance on what a fraud examiner should do based on what the analytics were telling them about their data. For example, if the analytics showed high risks for bribery and corruption in vendor payments, the software could recommend key steps, relevant company policies or guidance, sometimes even before making a payment in question. Theoretically, this approach could work, but the combinations were just too vast to anticipate every potential outcome. We needed more data. We got close with the "digital twin" concept with GE back in 2018, but we still couldn't acquire enough data to accurately prescribe each outcome. (See "'Profit & Loss-of-One': Preventing fraud, enhancing compliance using digital twins," by EY Fraud Investigation & Dispute Services and GE executives; Ed. Vincent M. Walden, CFE, CPA, *Fraud Magazine*, January/February 2018, tinyurl.com/273rpcj7.) What we were reaching for was prescriptive analytics. And unfortunately, it remained at the time a conceptual — rather than realistic — goal for compliance, fraud prevention and detection.

Gartner has long described four types of analytics that organizations use to drive decision-making from an analytics maturity perspective. (See Figure 1.) At the base, we have the hindsight of descriptive analytics, telling us what's already happened. Next, we have diagnostic analytics, which might tell us why something happened. We then move towards insight with predictive analytics, which shows us what will happen. Finally, we have the optimal, "pie-in-the-sky" vision of prescriptive analytics, which provides the foresight to implement or resolve something. (See Gartner.com.)

Figure 1



Source: Gartner

Descriptive and diagnostic analytics

Fraud examiners are quite familiar and comfortable with the first two types: descriptive and diagnostic analytics. They consist of traditional rules-based tests, computer-aided audit techniques (CAATS) and queries using spreadsheet and database tools that encompass the matching, querying, filtering and sorting we do to look for patterns and trends and control for weaknesses in the data. While easily used and rapidly deployable, the main difficulty of descriptive and diagnostic analytics is that they primarily focus on the past (which is good for investigations, expert witnesses or audits, but not when we seek to prevent and detect fraudulent activities). Once descriptive or diagnostic analytics have been applied, it's up to the fraud examiner to ask how or why those trends or anomalies occurred.

What pressure, opportunity or rationalization might be present outside the data that caused the breakdown in the first place, or what can be done in the future to avoid it?

Predictive analytics

Moving up the maturity curve, we have predictive analytics, which is just what it sounds like — it seeks to predict likely outcomes and make educated forecasts based on historical data. In eDiscovery, this technique was all the rage over the past 15 to 20 years when investigators could find a few "hot," "responsive" or "privileged" emails or documents and use predictive analytics (also known among litigation, investigation and eDiscovery professionals as "technology-assisted review") to find statistically similar documents. In a previous Innovation Update column, I demonstrated how to use predictive analytics with structured, transactional data to identify high-risk vendor payments during a U.S. Department of Justice investigation. (See "Using technology-assisted review to uncover suspicious transactions," *Fraud Magazine*, November/December 2022, tinyurl.com/r3ah4bv5.) Simply put, predictive analytics extends trends into the future to show possible outcomes. This is a more complex version of data analytics because it uses probabilities for predictions instead of interpreting existing facts.

Statistical modeling or machine learning is commonly used with predictive analytics. It might answer investigative questions such as whether your payments data include transactions statistically similar to those you've previously determined to be fraudulent. It's like saying "find me more like this."

The primary challenge with predictive analytics is that the insights it generates are limited to the data, and in a fraud risk management context, most companies typically don't have huge amounts of fraudulent transactions to train an effective model. This means that small or incomplete datasets won't yield predictions as accurate as large datasets might. In another recent Innovation Update column, I described some of my anti-corruption research out of MIT showing that when companies collaborate to share information about third-party payments and high-risk, potentially fraudulent transactions, they have a 25% greater chance of predicting improper payments than when each company's model is run in isolation. [See "From many, comes one (algorithm)," *Fraud Magazine*, March/April 2023, tinyurl.com/2dap499x.]

Prescriptive analytics

At the highest maturity level, we arrive at prescriptive analytics. Prescriptive analytics for fraud prevention and detection has been on my wish list for some time, but I've only had moderate success in achieving it. Here's why: Prescriptive analytics uses data from a variety of sources — statistics, machine learning, data mining — to identify possible future outcomes and show the best option while also suggesting what to do next. Prescriptive analytics is the most advanced (and difficult) of the four types of analytics because it provides actionable insights instead of just raw data. It enables you to envision future outcomes or risks and to understand why they will happen. In an ideal state, prescriptive analytics also can predict the effect of future decisions, including the ripple effects those decisions can have on different parts of the business. But let's not get too ahead of ourselves.

Generative AI and the use of large language models (LLM) that are trained on vast amounts of data to interpret and generate human-like textual output might just be the solution to get us to achieving prescriptive analytics for fraud prevention and detection.

OpenAI's ChatGPT, which incorporates a conversational chatbot with LLM to create content on the fly, is a great example of how an LLM using generative AI can be asked any question, and

Generative AI and the use of large language models (LLM) that are trained on vast amounts of data to interpret and generate human-like textual output might just be the solution to get us to achieving prescriptive analytics for fraud prevention and detection.

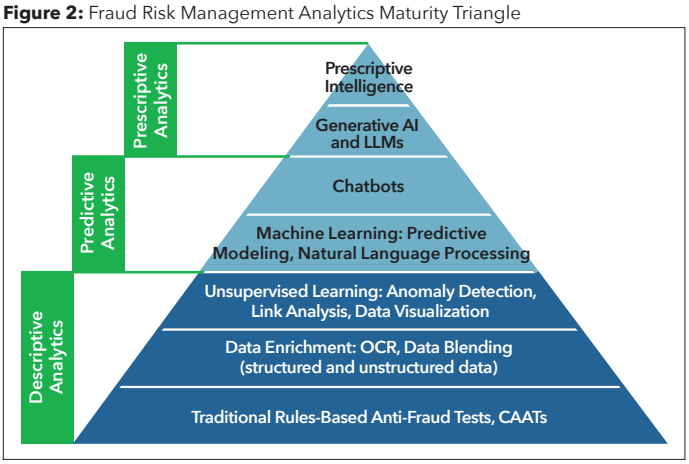
the response set is intelligent and most often quite useful — perhaps even prescriptive. LLMs don't have to be sourced only from the internet, like OpenAI, Google, Meta and the other large techs are doing. LLMs can also be domain-specific, such as fraud risk management, and can include information about your organization's training materials, policies, hotline data, contracts, ACFE training materials, U.S. Department of Justice and U.S. Securities and Exchange Commission enforcement information, and even your transactional data from vendors, customers or employees. Something big is about to happen — and I couldn't be more excited to enter 2024 with you as an anti-fraud professional.

The road map to prescriptive analytics

During the holidays last year, I had a vision that knocked me out of my chair and inspired this article. With all the hype around generative AI, machine learning and other such analytics techniques, it dawned on me that they all fit together like building blocks. Think of it like a triangle. (See Figure 2.) It's not as cool as the famous Cressey Fraud Triangle we all know and love, but it does put things in perspective. I call it the "Fraud Risk Analytics Maturity Triangle." At its foundation, you have your library of anti-fraud tests and controls. There are over a thousand of these tests, and my friend and mentor David Coderre helped spearhead many of them in his book, "Computer-Aided Fraud Prevention and Detection," which helped kickstart my career over 20 years ago. As you ascend the maturity scale, you see data enrichment. If you remember the hype around big data, you'll remember that everything was about blending structured and unstructured data that described high volumes, varieties (i.e., structured and unstructured data) and velocities of data within an organization. As part of that big data boom, unsupervised learning techniques such as anomaly detection, pattern and link analysis, natural language processing, data visualization (dashboards, for example) became popular. Keep in mind, all these techniques analyzed data in the past. Even today, most organizations' anti-fraud analytics capabilities are still descriptive in nature. They're the foundation of any good program. Hence, they're at the bottom of the pyramid, but there's more that can be done.

Now we get into predictive analytics. Let's start with machine learning, the core of predictive analytics we explored earlier in this article. But what about chatbots, which are simply computer programs designed to simulate human conversation? I put them into the predictive analytics category for now, given that there are

Figure 2: Fraud Risk Management Analytics Maturity Triangle



two types of chatbots: descriptive and predictive. The descriptive type is known as a "declarative chatbot," which works from scripted responses to hold rules-based or structured conversations with users. Think of them as interactive FAQs that can handle common questions. More interesting are the predictive chatbots, which are sophisticated, interactive and conversational. Also known as virtual or digital assistants, they use natural language processing, AI and machine learning to understand behavior patterns and user profiles. Familiar examples include Amazon's Alexa, Apple's Siri and Google Assistant. (See "What Is A Chatbot? Everything You Need To Know," by Shweta and Kelly Main, Forbes Advisor, Aug. 21, 2022, tinyurl.com/3yvz7kdr.) I include chatbots in my triangle because I think they'll soon supplement or perhaps even replace Tableau or PowerBI-type dashboards for use in fraud prevention and detection, as users will no longer need to click around in dashboards — they can simply ask questions of the data and the results and advice will be provided — perhaps in a dashboard or another format.

Finally, we get to the top of the pyramid and into prescriptive analytics. Remember, you can't have effective anti-fraud prescriptive analytics until you've completed many of the lower descriptive and predictive elements. Having a customized LLM tuned to your industry, your company's policies, your company's risk assessment and then combined with a decade of regulatory enforcement data, best practices guidance, fraud risk management training and even transactional data as previously mentioned could really drive intelligent insight if properly served to the end user at the right time and right context. I'm calling it "prescriptive intelligence," but it's prescriptive analytics at its best. The power and indexing capabilities of today's commercial LLMs are beginning to unlock intelligent responses on what users should do when presented with certain risks, patterns or anomalous events. It's not perfect yet, but I'm seeing it work well for several of my test customers. I'm predicting that 2024 will be the year for prescriptive analytics to flourish among companies.

Excitement for the future

Nikola Tesla famously opined that there isn't "any thrill that can go through the human heart like that felt by the inventor" as they see one of their creations coming to fruition, and "such emotions make a man forget food, sleep, friends, love, everything." As I plan my 2024 innovation roadmap around generative AI and prescriptive analytics goals, I can feel Tesla's enthusiasm, especially as I sit here on a Saturday night writing this column. Keep innovating! ■ **FM**

Vincent M. Walden, CFE, CPA, is the CEO of Kona AI, an AI-driven anti-fraud and compliance technology company providing easy-to-use, cost-effective third-party payment and transaction analytics software around corruption, investigations, fraud prevention, internal audit and compliance monitoring. He welcomes your feedback and ideas. Contact Walden at vwalden@konaai.com.