Overcoming data challenges in forensic investigations
The foundation for integrated human and machine intelligence

Traditional corporate antifraud measures are quickly losing ground against fraud schemes that continue to grow in both frequency and ingenuity. Internal and external perpetrators draw from a menu of ploys, including procurement fraud, employee expense fraud, financial statement fraud, bribery and asset misappropriation, such as intellectual property and data theft.

These threats alone provide impetus for companies to consider new approaches to antifraud and enterprise risk management programs. However, compliance pressures are raising the stakes even more. Regulators increasingly expect companies to have controls and monitoring in place to avert fraud-related issues involving FINRA guidelines, Foreign Corrupt Practices Act (FCPA) compliance, Sarbanes-Oxley requirements, and other dictates.

As discussed in an earlier Deloitte point of view, the evolution of fraud risk management and forensic investigations involves the application of analytics to transactions and data, using the insights derived from the integration of human and machine intelligence to refine and improve fraud-fighting efforts. Organizations across industries, and regulators themselves, are starting to use integrated, data-driven analytics approaches to identify potentially fraudulent transactions. Those that do not could rapidly fall behind and face increasing financial, reputational, legal, and regulatory risks.

One underlying factor that will weigh heavily in the value and effectiveness of analytics and monitoring activities is the data itself—how good it is, and how well it is used. Data can make or break analytics-driven forensic investigations.
Data challenges abound
An array of factors can contribute to gaps and shortcomings in monitoring fraud and conducting an investigation, including:

**Vast amounts of data.** Companies now electronically collect, process, and store more information than was imaginable even 10 years ago. And while the growth in data volume is impressive, even more so is the expanding variety of data sources generating that volume, including personal and business mobile devices, Internet of Things-connected devices, social media platforms... the list goes on and continually expands. Collecting, managing, monitoring, and analyzing the data that is most relevant to antifraud activities is already a complex process, and will only become more so.

**Inadequate capture and storage.** Legacy systems were often designed to capture information for a specific purpose, so the data available may not be rich enough for meaningful analytics. For example, transaction time stamps and the identities of employees performing transactions might not be captured. In some cases, too, only current data is available, historical information that is crucial for forensic analytics may not be stored. These problems may be exacerbated if the systems have not been updated regularly and additional information is not made available for analysis.

**Limited data accessibility.** A company with decentralized operations and data sources that are siloed by geographies and departments may lack a master system to consolidate data globally, inhibiting cross-correlation. Large global investigations can involve multiple countries, each using a different financial reporting or ERP system, making it more difficult to extract and analyze data. Jurisdictional data privacy and protection mandates can also limit access.

**Inadequate skillsets to process and analyze big data.** When data volumes are small, basic analytical skills and spreadsheet programs might be adequate to handle preliminary analysis of structured data from enterprise systems and other software applications, as well as unstructured data such as emails, texts, and voice recordings. But when that volume grows into the millions, analysis can require technology, advanced analytics, and forensic skills that aren’t readily available within many organizations. The technology and training investment required to support next-level fraud monitoring can be substantial.

**Static reporting designed for business as usual.** Legal, compliance, and internal audit teams may encounter barriers to gathering data from sources such as the finance, IT, procurement, and sales departments. Standard reports from those groups may provide only limited information; for example, in the context of procurement, identifying information such as a vendor contact name, address, and phone number might not exist in a standard vendor report, which could limit the ability to compare vendor contact information to employee data to determine potential overlap. Often, when reports were designed, parameters were established poorly. Or they may have been created years ago when the types of information investigators or compliance might need today weren’t even considered.

**Lack of diverse data to correlate findings.** Companies may not adequately explore external data sources, such as third-party reporting databases and social media, to capture a comprehensive view of the fraud risk presented by a company’s supply chain, sales channel, and employees.

Any one of these challenges by itself can slow a legal or compliance team’s efforts to apply machine learning and cognitive analytics. Together they represent a significant barrier if they aren’t addressed in the ramp-up to using advanced artificial intelligence capabilities to better identify and deter fraud.
Data considerations for analytics-driven fraud risk management

Organizations can take several steps to prepare an effective foundation for analytics-driven investigations and fraud monitoring:

**Involve stakeholders in building the transformation roadmap.** Specific areas of a company may be primed and ready to undertake analytics-driven fraud risk management, but others need in on the plans, too. Internal audit, legal, compliance, IT, and the businesses can all have roles and interest in the analytics efforts. Discussions with relevant stakeholders can identify synergies and ways to leverage technologies in use elsewhere in the organization. And, stakeholders can help identify high-risk areas that warrant focus, such as time and expense reporting, vendor management, and third-party payments (see “Choosing a starting point”). Also, by keeping in contact throughout the analytics initiative, data scientists can stay aware of evolving business needs and business users can understand how data is being stored, accessed, and secured.

**Centralize as much data as possible to support fraud monitoring.** While centralizing all enterprise data would be the Holy Grail for the fight against fraud, it may not be realistic in many organizations today due to disparate data sources, geographic locations, and gaps in systems integration. Still, emphasis should be placed on bringing as much data together as possible to maintain data integrity, consistency, and control and for enhanced fraud monitoring, analysis, and insights. A good starting point is consideration of requirements for and possible impediments to drawing data from different departments and geographic regions.

**Establish secure, structured access to data.** A compliance department planning to conduct analytics can benefit by defining early on how data will be handled, where it will be stored, and who will be allowed access to it. Considerations include needed safeguards against breaches and policies and procedures for treatment of personally identifiable information (PII) and other sensitive data.

**Incorporate relevant external data.** External data can be brought into the centralized repository to cross-correlate with internal data.

**Begin to lay a solid technology foundation.** It is important to plan for investment in technology and software applications, as part of an overall enterprise solution, that can support effective data collection and analysis for fraud monitoring and to leverage the same data for multiple purposes. The technology should be scalable so both structured and unstructured enterprise data can be included in the analysis.

**Better data, richer forensic investigation, and fraud risk management**

The success of an analytics-driven fraud risk management program relies on the availability and accessibility of accurate, relevant, and rich data from different geographical locations, service lines, products, and external data sources. As mentioned previously, a centralized, enterprise-wide data solution would be optimal, but in its absence companies can still significantly improve their fraud monitoring and forensic investigation by considering these questions:

- What is the strategy to manage the ongoing proliferation of data?
- What type of analytic capabilities would fit the organization’s specific needs?
- Can tools or insights serve multiple purposes across the organization?
- What are key technology trends within the industry and how will the organization’s transformation roadmap keep the organization ahead of the industry?

The transformation to an analytics-driven program, including answers to these questions, is likely to require significant time and effort. As typical in the rollout of a new technology, a pilot program using a Test/Prove/Implement/Scale/Repeat methodology can be a helpful starting point. Focusing on early results while staying attuned to the big picture can help equip organizations to address future fraud risks.

**Choosing a starting point**

Ask a risk and compliance professional to identify fraud risks that would be top candidates for advanced analytics techniques, such as machine learning and cognitive computing, and you may well hear about dozens. One risk team, knowing it would have to show return on its analytics investment to secure funding for broad deployment, distilled down its list of over 100 areas and chose three in which to begin the analysis. The demonstrated value of these initiatives supported expanding the analytics effort to additional risks. The lessons learned: Start small, pick smart, drive value.