Portfolio Media. Inc. | 111 West 19th Street, 5th Floor | New York, NY 10011 | www.law360.com
Phone: +1 646 783 7100 | Fax: +1 646 783 7161 | customerservice@law360.com

The Rise Of AI In Corporate Investigations

By **Zachary Adams and Richard Chalk**
September 6, 2017, 12:43 PM EDT

Zachary Adams

Richard Chalk

The recent growth of corporate data — both in volume and complexity — is exponential and unabated. A 2016 report from IBMestimated that 90 percent of the world's data had been generated in the previous two years. With new systems and devices constantly being developed, this trend is sure to continue. One byproduct of this growth has been a commensurate increase in the complexity of corporate investigations, as lawyers are faced with vast amounts of data to sift through. In response, artificial intelligence is increasingly being incorporated into government investigations to organize and analyze large, diverse sources of data and documents.

AI has made inroads at both the corporate and government level. In-house counsel and compliance teams are now relying on AI technology prophylactically to identify risks before they escalate. In response to government investigations, larger companies have been utilizing limited to advanced forms of AI technology for a number of years. Similarly, the U.S. Department of Justice, the U.K. Financial Conduct Authority and other government enforcement authorities have begun to explore and deploy advanced data analytic tools. As regulators implement AI technology for their own enforcement and supervision efforts, they will naturally become more comfortable with companies utilizing the technology as part of their own internal investigations.

Used correctly, AI will drive efficiency in investigations, maximize results and help to promote regulatory compliance.

Unstructured and Structured Data as the Foundation of AI

The breadth and scope of corporate data being generated has created numerous challenges for investigators seeking information in criminal, civil and regulatory cases. Further complicating matters, the explosion of information being generated by disparate sources has forced companies to transition their data to repositories that are often less well organized, including data lakes that store native or raw-form data.

AI can focus large-scale reviews of emails and other documents, i.e., "unstructured data," an essential

component of many internal investigation and the area AI has gained the most industry traction. AI can be used to search all sources of information, from mobile device apps to scanned hard copy documents, to locate and escalate relevant data for review. AI looks for keywords and patterns in data, evaluates the information and collates a set of results. When deployed effectively, AI can substantially reduce the universe of data for review, allowing professionals to hone in on the most relevant material.

The most effective AI investigations couple continuous active learning technology with concept clustering to discover the most relevant data in documents, emails, text and other sources. Continuous active learning is a form of supervised machine learning which allows AI to recognize keywords and patterns in data, learn and adapt from the information and continuously score and rank data for its relevance. Concept clustering uses unsupervised machine learning algorithms to examine large batches of data, identifying contextual similarities and creating logical groupings. Supervised machine learning AI requires human input to drive learning, while unsupervised generates insights without any input.

AI can also be an effective tool for investigations that require synthesizing structured data, such as information contained in a database. The technology can be tuned to efficiently integrate and query massive data sets to eliminate noise, connect seemingly unrelated data points and identify key patterns. As but one example, AI was recently used to search for patterns in years of domestic flight data, ultimately identifying a prevalence of spy planes being used by federal and state actors. A similar approach can be used to scrutinize financial transaction databases for evidence of money laundering.

Additionally, analytics can incorporate findings from the unstructured review into its analysis of the structured data. If, for example, a particular customer account of note is identified by a lawyer in an email, the technology can be tuned to identify other similar customers based on multiple data features. For example, those with a similar address or those who were on-boarded in a similar timeframe by the same employee. Integrating the unstructured and structured work streams, which requires close coordination between lawyers and the technical data team, is the most effective form of supervised machine learning.

This technology is applicable to countless types of internal investigations and is particularly relevant to investigations with a financial crime component, including fraud, insider trading and violations of economic sanctions. When the unstructured and structured analytics are used in tandem, AI technology can increase efficiency by minimizing cost and maximizing output. As AI continues to evolve and improve over the lifespan of the project, the benefits of an effectively implemented approach will also increase.

The Uptake of Government and Regulatory Usage of AI

The use of AI technology by enforcement authorities is nothing new. In fact, in 1986, then-Attorney General Edwin Meese III provided a statement to the House Committee on the Judiciary in which he discussed budgetary needs for AI technology. He noted broad applicability for the technology in the context of complex criminal investigations involving narcotics, terrorism and organized crime, and also foretold of “a number of important new and expanded uses for automated data processing planned within the Department.”

More recently, enforcement authorities across the world are ramping up their own use of AI technology for enforcement and supervisory monitoring. Just last year, the Financial Conduct Authority in the U.K. announced that they were seeking new data science tools to help further its industry oversight. The U.K. Serious Crime Office has been utilizing AI to spot financial crime and corruption. The U.S. Securities and Exchange Commission has similarly used AI and behavioral analytics to augment its enforcement activity.

Last year, the White House Office of Science and Technology Policy (OSTP) held a workshop describing the many uses for AI, which included uses in the criminal justice context. Releasing a report in October last year, "Preparing for the Future of Artificial Intelligence," the National Science and Technology Council, which falls under the OSTP, repeatedly pointed to criminal justice as an area of increased AI usage. They noted:

AI also has the potential to improve aspects of the criminal justice system, including crime reporting, policing, bail, sentencing, and parole decisions. The Administration is exploring how AI can responsibly benefit current initiatives such as Data Driven Justice and the Police Data Initiative that seek to provide law enforcement and the public with data that can better inform decision-making in the criminal justice system.

The focus of the OSTP on AI ultimately led to the creation of the Subcommittee on Machine Learning and Artificial Intelligence, which was established under the National Science and Technology Council to advise, coordinate and support initiatives involving AI. Not surprisingly, the DOJ is represented on this subcommittee.

This government initiative has already produced tangible results. In August 2017, the DOJ announced a task force targeting opioid abuse, the centerpiece of which is a data analytics program monitoring opioid prescriptions nationwide. According to Attorney General Jeff Sessions, the data analytics could assess, among other things, "which physicians are writing opioid prescriptions at a rate that far exceeds their peers" and "pharmacies that are dispensing disproportionately large amounts of opioids."

Treading Carefully When Rolling Out AI for Internal Investigations

While many regulators utilize AI themselves, they will want to ensure that corporations relying on AI as part of an internal investigations are being as thorough and diligent as possible. The regulator will need to understand why AI makes sense in the instant matter; the technology that is being employed; the various sources of data that were incorporated; and the oversight, supervision and quality control processes in place. It is essential to have an open and constructive dialogue with enforcement authorities regarding the chosen methodology. One strategy for highlighting the reliability of the AI methodology is to demonstrate its effectiveness on a small sample of data and then extrapolating that to a much larger population.

Corporations and their counsel should be mindful of the fact that the government often has access to multiple sources of documents and data during the course of an investigation and that this information can be used to independently verify the quality of the AI methodology. These other sources could include counterparties to a transaction, financial institutions and individual witnesses.

Another contentious topic is the ability to train AI, based on the experiences of multiple internal investigations. While this may be possible where the investigations occur at the same company, ethical and legal challenges arise when commingling data and learnings from different companies. For example, local banking secrecy or data privacy laws often place substantial restrictions on cross-border data transfers. As AI relies on underlying data which may be subject to these restrictions, the technical data team must intimately understand the data being ingested and incorporated into the algorithm. The internal learnings acquired by human investigators are not subject to transfer restrictions for obvious reasons.

It is essential that the technical data, legal and compliance teams work together when integrating AI into an internal investigation. Each company will have its own method of storing data, differentiating legacy data and linking data that originates from different systems. Every investigation will have its own set of facts and pertinent legal issues. Only a bespoke approach will allow AI technology to find relevant data in the most efficient and cost-effective manner, while also conforming to a company's policies and legal requirements. The increased use of AI technology may be seen as daunting, but the explosion of corporate data only amplifies the need for more efficient corporate investigation tools.

Zachary Adams is a partner in the Washington, D.C., office of Squire Patton Boggs LLP. Richard Chalk is a director in the London office of Navigant Consulting Inc.