INCREASED USE OF BIG DATA IN SEC ENFORCEMENT

The SEC's new analytic tools, applied to massive data sets, have enabled the Commission to proactively identify and pursue a significantly increased number of enforcement actions in the last few years. Former chair Mary Jo White has praised the new methods as "transformative." The authors describe this development, the various data analytic programs the Commission now uses, and some recent cases brought as a result. They close with practice tips for counsel.

By Elizabeth P. Gray and Catherine E. Fata *

Since the establishment of the Securities and Exchange Committee in 1934, its mission has remained constant — to protect investors and, derivatively, the financial markets through regulation and, over time, enforcement of the federal securities laws. The SEC's methodologies for identifying and investigating financial misconduct crucial to achieving its mission have evolved over time. Until relatively recently, the SEC's Division of Enforcement generally investigated misconduct referred to them by third parties both within the agency, such as operating divisions, and outside of the agency, such as FINRA. But the introduction of "big data" has dramatically changed the way business is done at the SEC. By successfully harnessing the power of big data, the SEC has transformed its enforcement program to one in which it more proactively identifies problematic conduct and tests its market theories. The SEC now has at its fingertips the technology to rival the most sophisticated companies in the world. This has dramatically changed the way the SEC operates, and companies and corporations subject to their jurisdiction must take notice.

THE DAWNING OF A NEW ERA: INCREASED USE OF "BIG DATA"

Historically, the SEC's Division of Enforcement relied heavily on third parties as a source of investigations. Traditional sources of referrals were financial disclosures, whistleblowers, and self-regulatory organizations, such as FINRA. As recently as five years ago, the SEC was originating far fewer investigations, but new technology — such as the Corporate Issuer Risk Assessment (CIRA) and technology within the Division of Enforcement's Analysis and Detection Center — has allowed the SEC to take an increasingly proactive approach to enforcement. Former SEC Chair Mary Jo White has lauded this new, data-driven approach as "transformative." By frequently originating their own investigations, the SEC is not only less reliant on

---

1 Mary Jo White, Chair, SEC, Speech at the New York University School of Law Program on Corporate Compliance and Enforcement: A New Model for SEC Enforcement: Producing Bold and Unrelenting Results (Nov. 18, 2016).

2 Id.

---

*ELIZABETH P.GRAY is a Partner and CATHERINE E. FATA is an Associate at Willkie Farr & Gallagher LLP. Their e-mail addresses are and cfata@willkie.com.
traditional sources, but it also has drastically increased the number of enforcement actions brought in each successive year. For example, in 2016, the SEC brought 868 enforcement actions, many of which originated in-house.\(^3\) By contrast, the SEC brought just 630 and 574 enforcement actions in 2005 and 2006, respectively.\(^4\)

The impact of the SEC’s reliance on big data has allowed the Commission to adopt a much more proactive approach to investigations. In contrast to traditional referral sources, the use of big data gives the SEC more control over the companies and conduct it chooses to investigate. The proactive approach can also be seen in the investigation itself. Where the Enforcement Division was previously reliant on cooperation from registrants, it now has more control over fact-finding by testing hypotheses and data throughout the enforcement investigation. Cooperation remains an essential aspect of the SEC enforcement model, but the Commission can now use its own data as a check on the version of the facts that the company relays. Internal investigations aren’t going anywhere, but more than ever, the SEC has the data and technology necessary to be an active participant. The importance of this proactive approach can be seen in the allocation of significant SEC resources to data analytics, the types of investigations that are conducted, and the enforcement actions that are ultimately brought.

**WHAT IS “BIG DATA”?**

Scott W. Baugess, Deputy Director and Deputy Chief Economist at the Division of Economic and Risk Analysis (“DERA”) at the SEC, has said that big data is best defined by a rule of thumb: anything larger than the amount of data that fits on a thumb drive is “big data.”\(^5\) He elaborates that “big” data is anything that can’t be processed in a standard desktop application such as Microsoft Excel.\(^6\) An alternative view is that big data is anything that takes longer than a day to process, or data that takes longer to process than it took to create (i.e., if it takes longer than a day to process a day’s worth of trading data). However, “big data” cannot be defined by size alone. As Baugess has pointed out, because of continuing advances in computing, what was “big” last year is not big today.\(^7\) There can be no static definition, as the understanding of big data continues to evolve as technology progresses. Thus, big data is best understood as any data that approaches current computational limitations on analysis.

The main source of data for the SEC is information provided as part of mandatory public disclosures. The SEC regularly receives massive amounts of data from registrants, far more than can be analyzed by individual reviewers. Though the SEC has always had access to registrant’s disclosures, advancements in computer technology now allow much more data to be collected and stored. The data has been there all along, but now more than ever, the SEC can effectively mine this data for information, trends, and other patterns.

The creation of DERA in 2009 highlights the emphasis that the SEC places on developing data analytics technology. Considered the “think tank” of the agency, DERA was created to “integrate financial economics and rigorous data analytics into the core mission of the SEC.”\(^8\) DERA employs attorneys, economists, analysts, statisticians, and computer programmers — all charged with developing and providing economic and statistical analyses that are utilized throughout the agency. DERA is also responsible for developing analytical approaches, methods, and models used to identify issues, trends, risks, and potential securities violations.\(^9\) For example, DERA developed a broker-dealer risk assessment that

\(^3\) Id.


\(^5\) Scott W. Baugess, Deputy Director and Deputy Chief Economist, Division of Economic and Risk Analysis, SEC, Speech at Midwest Regional Meeting of American Accounting Association: Has Big Data Made Us Lazy? (Oct. 21, 2016).

\(^6\) Id.

\(^7\) Id.

\(^8\) Economic and Risk Analysis, SEC, https://www.sec.gov/dera/about (last accessed May 18, 2017).

\(^9\) Id.
helps SEC examiners allocate resources by assessing a broker-dealer's riskiness relative to its peer group by comparing metrics such as operations, financing, workforce, firm structure, and type of dealing activities.\textsuperscript{10} DERA has been considered a great success, proven by the fact that it now encompasses 10 offices, including the Office of Corporate Finance, the Office of Litigation Economics, the Office of Research and Data Service, and most recently, the Office of Risk Assessment.\textsuperscript{11}

THE CORPORATE ISSUER RISK ASSESSMENT

In 2014, the SEC created the Office of Risk Assessment, a division within DERA that focuses on financial and risk modeling related to corporate issuers, broker-dealers, investment advisers, exchanges, and trading platforms.\textsuperscript{12} The most important programs administered by the Office of Risk Assessment combine DERA's analytic and financial modeling capabilities with the institutional experience of agency attorneys. One such program is the Corporate Issuer Risk Assessment (CIRA), which expands on an accounting model to detect anomalous patterns in financial reporting that may warrant additional inquiry. CIRA aggregates and organizes the massive amounts of data the SEC regularly receives from registrants.\textsuperscript{13} CIRA generates lists of companies or their filings that meet criteria for further analysis and compares filings across different companies to detect abnormalities.\textsuperscript{14} Additionally, SEC staff attorneys have access to CIRA by way of a user-friendly Microstrategy dashboard platform. The dashboard platform incorporates over 100 custom metrics that attorneys can search and analyze. For example, an SEC attorney might be interested in how inventory at a company is moving relative to sales, because inventory buildup might lead to more aggressive accounting.

As Baugess recently told reporters, "[CIRA] basically serves as instant corroboration to thoughts or ideas they may have." If a tip or referral comes in on a particular issuer, SEC staff attorneys can utilize CIRA to identify indicators of earnings management or other red flags corroborating the complaint.\textsuperscript{15}

When it was first announced, CIRA was criticized as a financial "RoboCop." However, CIRA has become an integral part of the SEC's efforts and has also led to the development of more comprehensive detection programs. On November 15, 2016, the SEC announced it had approved Rule 613, a plan to establish a national market system to create a single, comprehensive database, or "consolidated audit trail" ("CAT") — a database that stores every trade issue, execution, and cancellation.\textsuperscript{16} Currently, the SEC must obtain and merge together large volumes of disparate data from different entities when analyzing market activity.\textsuperscript{17} The CAT will be the first time that the SEC has access to a single database of readily available, comprehensive data regarding trade orders and executions. The plan, which was developed by FINRA and other national securities associations, will be the world's largest data repository of securities transactions.\textsuperscript{18} It is predicted to ingest 58 billion trade events on a daily basis.\textsuperscript{19} Under the plan, regulatory organizations will be required to begin reporting to the CAT within one year of approval, with large broker-dealers following within two years of approval and small broker-dealers within three years.\textsuperscript{20}

One key impact of the CAT will be an increased ability of the SEC to recreate previous market conduct in any investigation or enforcement action. Presumably, SEC attorneys could use the CAT to run analytics for any public company on essentially any metric of interest. For example, the SEC may announce that it will be taking an increased focus on non-GAAP disclosures. An SEC attorney may use the CAT to run analytics for any


11 See supra note 8.

12 Id.

13 See supra note 10.

14 Id.


19 Id.

20 Id.
language that is consistent with non-GAAP disclosures, and compare the findings to other indicia of previous problems. This is but one example of the almost limitless ways that the SEC can apply its institutional priorities and knowledge to the vast amount of data available to it.

**ANALYTICS WITHIN THE ENFORCEMENT DIVISION**

The SEC’s Enforcement Division has developed its own data analytic capabilities with a staff that collaborates with DERA. The Center for Risk and Quantitative Analytics (CRQA) was established in July 2013 to support and coordinate the Division’s risk identification, risk assessment, and data analytic activities by identifying risks and threats that could harm investors, and assist staff nationwide in conducting risk-based investigations and developing methods of monitoring for signs of possible wrongdoing. It works in close association with other Commission offices and divisions, especially DERA, and provides guidance to the Enforcement Division’s leadership on how to allocate resources strategically in light of identified risks. It has been used to investigate and build cases against more than 200 entities and individuals.

"More than ever, the SEC is developing in-house innovative analytical tools to take advantage of today’s data-rich environment," said White at the end of her tenure as Chair. "The result is that the number of cases we are able to originate in-house has risen dramatically. In each of the last three years, the SEC has broken its previous record for the number of cases brought," she said.

Recent SEC enforcement actions provide insight as to how the SEC can effectively use big data. One of the Division of Enforcement specialized units, the Market Abuse Unit, used its internal Analysis and Detection Center, founded in 2011, to bring at least nine insider trading cases over the past few years from leads generated solely from data analytics. The SEC used data analytics to search records of individuals who had made well-timed trades ahead of corporate news, and in doing so, exposed insider trading schemes. In one such case, Steven McClatchey, a former investment banker at Barclays PLC, provided nonpublic information about client transactions to his friend Gary Pusey. Pusey then traded on this information. The illicit trades, which the SEC’s Market Abuse Unit says it caught through its Analysis and Detection Center, netted Pusey $76,000. Both McClatchey and Pusey pled guilty in the Southern District of New York in July 2016.

Recent cases also indicate a renewed interest in financial reporting and related technology. Margaret McGuire, the chair of the SEC Enforcement Division’s Financial Reporting and Audit Group, which was formed in 2013 with the stated goal of strengthening the agency’s efforts to identify and prosecute securities law violations related to financial reporting and audit failures, has repeatedly referred to CIRA as an important source of investigations. She has also referred to the Issuer Monitoring initiative by which the Commission uses a methodology to review the data of certain issuers and has identified 250 issuers that are of interest to the SEC. While the SEC does not regularly identify the use of CIRA in announcing financial fraud actions, recent SEC cases indicate a renewed interest in financial accounting. In June 2015, for example, the SEC reached a $190 million settlement related to allegations of falsified public disclosures by Computer Sciences Corporations ("CSC"). Five executives also paid fines. Allegedly, CSC engaged in manipulation of financial results and failed to disclose that the company would likely experience material adverse financial consequences due to its failure to perform its most valuable contract.

Development of the CAT and other tools to analyze data will allow the SEC to efficiently focus on those issues it has chosen to prioritize. For example, the SEC could utilize big data to target gatekeepers by running

---

22 Mary Jo White, *see supra*, note 1.
23 *Id.*
24 *Id.*
25 *Id.*

June 21, 2017
analytics for all financial statements prepared by a specific accounting firm.

PRACTICE TIPS

Big data is an important tool for the SEC and should be used to increase the Commission’s efficiency. The SEC’s continued use of big data, however, raises new issues for public companies and financial service firms, as well as their counsel. The introduction of CIRA and the new CAT provides the SEC with a different perspective from its registrants. Prior to the advent of Big Data, the client had the informational advantage — but now the SEC can observe a company’s public statements against the backdrop of all other public filings. From this vantage point, the SEC may have insight into issues that the client may not, in good faith, even be aware of yet. Counsel should take the time to assess a client’s regulatory, disclosure, and accounting practices, and realize that clients may or may not be aware of the potential issues or vulnerabilities the SEC’s advanced technology may expose. It is also noteworthy that the SEC is able to identify insider trading patterns more readily through the SEC’s Analysis and Detection Center. Finally, in conducting internal investigations, counsel should be aware that the SEC may be able to test facts and conclusions through the use of its internal technology.