CORPORATE COMPLIANCE INSIGHTS
THE PREMIER GLOBAL NEWS SOURCE FOR COMPLIANCE, ETHICS, AUDIT & RISK

Getting to the Bottom of Whistleblower Complaints

by Jeffrey Grobart · Jan. 7, 2020

Whistleblower complaints are on the rise and may gain further traction from recent stories in the news, increasing the burden on the investigators. H5's Jeffrey Grobart discusses tools and methods to accelerate the vetting process.

From Edward Snowden to the as-yet-anonymous White House whistleblower(s), actions taken by individuals to expose suspected wrongdoing are making headlines, casting the whistleblower process into the spotlight.

These more notorious cases are just the tip of the iceberg.

According to the 2019 SEC Annual Report to Congress on the Whistleblower Program, which provides incentives and protection to whistleblowers under Section 21F of Dodd-Frank, the Commission received its second-largest number of whistleblower tips — more than 5,200, which represents a 74 percent increase since the whistleblower data began being collected in 2012. But whistleblower complaints go well beyond Wall Street.

The Occupational Safety and Health Administration (OSHA), the government agency responsible for protecting workers in a wide range of industries, reports a 29 percent increase in whistleblower complaints filed over the last five years, up to 9,566 in 2018, creating a significant backlog in light of a reduction in the number of investigators. More than 3,000 of those complaints resulted in a full investigation.

Electronic Information on the Front Lines

For any whistleblower complaints that do result in an investigation, the nature of the allegations will usually determine what steps will be taken to gather information required to pursue the claim. Today, almost any corporate investigation will involve scrutiny of an abundance of electronic information, triggering a review workflow not unlike a discovery effort for a litigation.

But investigations are different from litigation. Instead of looking for all of the relevant documents that relate to a known set of facts, the quest may be to find the key documents that can identify or establish a fact pattern in the first place, or — more difficult — to show there is no evidence that any fact pattern exists to support the claim. A small number of documents may actually tell the story — if only they can be found. That is where advanced tools, coupled with the right expertise, can play an important role.

The use of more advanced analytics and AI tools to target information is quickly gaining ground for both litigation and investigation efforts, but there is no one-size-fits-all solution. Rather, a combination of tools and methods should be considered, depending on the situation.

Accelerating a Response: Act Fast, and Begin with the Smallest Subset of Information

Like all investigations, speed is key in vetting whistleblower complaints. The challenge is to move quickly to determine if a broader investigation is warranted; dragging out an inquiry response just raises red flags and can be misconstrued as culpability.

Intelligently prioritizing high-value personnel sources and data stores enables a quick start, allowing an investigation to expand in an incremental and controlled way, branching out in various directions only as more information comes to light. A sound methodology involving data sampling and advanced search technologies to identify critical documents from a defined subset of targeted data can provide important information to guide further actions.

Data sampling (provided it is done correctly) is useful to provide insight into the larger data population, enabling development of a targeted effort to locate relevant information without wasting time on false leads. As relevant facts are exposed, the data pool may need to expand, but informed decisions based on strategic, measured assessments help contain data volumes under scrutiny and make the inquiry more efficient and cost-effective. This requires methods and expertise that go beyond the usual keyword search tactics to develop an iterative process that pays attention to the indexing and search capabilities of the search platform being used.

If relevant information cannot be found, well, good news! However, this will likely require scrupulous and convincing documentation of all steps taken to locate relevant evidence in order to demonstrate a level of thoroughness that will satisfy regulators. On the other hand, if compelling evidence is found to support the allegations, forensic analyses may be necessary to confirm legitimacy of certain data and ensure, for example, that material was not falsely created to support a claim or that incriminating data has not been deleted by bad actors.

Leverage Advanced Search, Analytics And AI

Today's advanced search, data analytics and artificial intelligence tools can help to quickly pinpoint meaningful documents, assuming there are any. The multidimensional character of electronic data can be leveraged to provide a variety of investigatory clues — from timing to geography — in addition to the exploration of textual content. For example:

- Analytic tools that can access and search metadata along with body text, headers and footers, can be useful in helping to prioritize data.
- Techniques such as setting index parameters, virtual deduplication and email chain analysis can reduce and enhance the data population, narrowing and enriching the data being interrogated.
- Analytic tools coupled with AI can be used to explore communication patterns, enhanced with dates, times and density of communications, that may reveal potentially "interesting" relationships.
- AI in the form of machine learning (see tools, below) can group like documents for faster review.
- Precision-focused searches by those with linguistic expertise can target particular language use, locating important documents early that can shorten timelines.

Know Your Tech Tools (and Your Experts)

There are various tools and methods that can be deployed to find critical information, and it is important to know which are most effective in a given circumstance. Certain technology-assisted review (TAR) tools require a robust data collection to be most effective, which is more useful in a litigation review effort than an investigation where a quick start and a targeted outcome is important.

Cluster Analysis
Leveraging cluster analysis can be an effective first step in understanding target data sets. Looking at how documents group together based on their content can make discarding uninteresting content a simple exercise. Likewise, if particularly relevant clusters are found, the underlying documents can then be prioritized for further scrutiny.

Continuous Active Learning (CAL)
When looking to review document sets too large to be reasonably assessed in their entirety in a reasonable amount of time, continuous active learning tools (CAL) can be very effective. This is especially true when data volumes and responsive content are moving targets. As an investigator reviews and categorizes documents, the CAL algorithm re-prioritizes the remaining population, enabling documents that are most likely to be of interest to be promoted to the top of the queue. Introducing new data to the investigation is not problematic since the algorithm will find and prioritize new documents of interest as they are added to the model.

Human Expertise
In addition to analytics and AI tools, human expertise is a prime component of any investigatory process. As mentioned earlier, linguistic search expertise paired with advanced analytics and search tools is a powerful combination. Based on topical information about the nature of the investigation, linguists can factor in expert knowledge of the ways in which human beings actually express themselves to create sophisticated search queries. After all, it could be not the presence, but the absence of certain language that may be revelatory, or the unusual recurrence of certain analogies (why are x and y always talking about weather reports?). Linguists can provide the groundwork for the most robust analytics.

Proactive Monitoring: The Best Approach

The best way to contain whistleblower complaints is to prevent reasons for them, but the ability to assess risk proactively is the next best thing. Identifying potential areas of risk is an important element.

If whistleblower activity or a certain type of bad behavior is becoming more common in an organization's industry sector, for example, that particular subject matter could be targeted for routine monitoring using AI and linguistic and analytics tools trained to flag worrisome language or sentiment in data stores. Health care insurance investigations, for example, are frequently on the front lines of whistleblowing activity and face challenging barriers in targeting massive and complex data stores that may require interrogation for any number of types of fraud.

Much like AI systems used to monitor and flag unusual financial transactions, custom classifiers can be deployed to proactively monitor unstructured data, such as email or document repositories, to identify correspondence or other information that contains content that should be scrutinized by in-house counsel or compliance officers. This approach goes beyond typical rudimentary email filtering options that are often used to flag questionable communications.

Early self-reporting of inappropriate actions generally leads to reduced penalties and should act as an incentive that counterbalances those that a whistleblower may find compelling. Overall, the best intelligence to apply in whistleblower actions comes not from artificial means, but from humans who provide standards, policies and ongoing scrutiny of corporate data and behavior that leave whistleblowers with no tune to sing.

Jeffrey Grobart is Associate Director of the Professional Services Group at H5, where he leads expert teams in pairing innovative technology with accumulated expertise to provide best-in-class service to corporate and law firm clients.

Jeffrey is an expert in the use of litigation technology across the EDRM and has provided consulting services in a wide variety of practice areas during his 20-year career, with an emphasis on anti-trust, pharmaceutical, commercial real estate and intellectual property engagements. He has led cross-functional teams in addressing a wide variety of e-discovery issues and has been a pioneer in helping deploy the use of predictive technologies in e-discovery and investigations efforts.

H5