Forex trading institutions tap eDiscovery technologies in their own monitoring workflows

By Duncan Gardiner & Gareth Meatyard / Epiq Systems

Epiq Systems is well known for its capabilities in audio review within the eDiscovery sector. Its specialities are regularly approached by financial institutions to discuss effective solutions in the monitoring space. Following are responses to some of the questions we hear most frequently:

Q: How do compliance teams in financial institutions deal with ongoing monitoring requirements?

During discussions with compliance teams, we have noticed that their monitoring workflows tend to be either non-existent, very restricted or unevenly applied. Monitoring tends to be applied in a much more vigorous way in relation to only a few refined data sets (e.g., email and possibly instant messaging). Epiq is often brought into a discussion initially to solve an audio problem, but soon becomes involved in the wider monitoring program and the goals of that organization, helping the organization build a monitoring workflow encompassing a wider data set, and reviewing processes and technology that can help.

Q: Where is "hot" information found?

Epiq partners with audio processing specialists in eDiscovery and also leverages these technologies in the monitoring space. One such partner, Intelligent Voice®, has carried out research into the typical responsiveness of target information. They suggest that while, a number of years ago, target conversations of wrongdoers were to be found in mainstream communications such as email (e.g., many of the evidentially significant emails contained in a recorded state because of deal room open mics or "squawk boxes") or phone calls. Audio evidence is a much more valuable data set to consider during an investigation, but it is also one of the more challenging forms of data to extract because the content is not in the same format as the written word, which forms the vast majority of the disclosure. Standard search terms used to find the written word will not always find the spoken word. Further, the spoken word is made up of sounds which can be very specific to the actual speaker, meaning that a search that works for one custodian may not work on another for exactly the same words in a sentence.

Q: Why do compliance teams need to implement these technologies?

Compliance teams need to implement these technologies in a variety of ways. The tools are designed to help compliance teams

Keeping current: Compliance teams must keep up with the media that companies and their employees have available to them to do business. For example, mobile technology is changing almost monthly. Applications are written by non-vendors and hardware is being created by a wide range of producers.

Maintaining access: Compliance teams need to establish and maintain access to those current communication media. The business will have created certain data universes that it controls and delivers. But if an organization has loose IT policies, such an organization may find that its employee base is using applications outside of its control, like Skype, Viber, WhatsApp. If the business has not put any controls around this usage, there may be conversations happening about financial transactions that are outside of the business

Volume: The increased number of communication methods brings with it an increase in the amount of information produced. The challenge then is how much of this information organizations need to monitor, and whether they have the technologies in-house to capture the relevant information, index it and apply the filters needed in order to properly review it.

Search: Filters and automated searching can remove much of the human element from the monitoring process.

Keeping what you don’t know: This process can only detect what you know about because it’s not doing anything ‘intelligent’ with the universe of data. It is simply applying the filters that the compliance team provides. New, unknown breaches will not be searched for unless an investigative capability is added to the monitoring workflow.

The right team: Monitoring teams need to consist of people with varying disciplines including those who understand the relevant legislation and others who understand the business and the way that its traders communicate. A team like this will be able to implement more effective testing of the wider data universe and undertake broader searches to identify new threats and patterns of behavior. A well-informed team will then take that knowledge and undertake investigations about any problems that they find, or if they find information that needs to be fed back into the automated system, they will manage the filters in the automated process to keep them current and up-to-date, ensuring that what is produced is more relevant, more fruitful and more targeted to what the investigators are looking for.

Q: What is the problem with silence?

Silence removal is helpful when dealing with audio, especially with recordings from squawk boxes or open mics because those recordings can be quite long (typically 8, 9, 12 hours long). During very long recordings, there may be only short periods when key or any conversations are taking place at all. If you remove the silence, you can reduce the reviewable content down to a much smaller volume of information, which reduces the overall cost of having to search and consider that specific recording.

Q: What other information should be harvested?

Duplication and metadata analysis should also be harvested. Many telephony systems create metadata information about the recording: telephone numbers, dates, times, names, participants in IM chats. All of that information should be harvested because it may be the first step to sorting, grouping and making a decision about whether to give further and deeper analysis.

Q: What is profiling in audio review? How does it help during the monitoring process?

Profiling in audio solves a very particular problem. Metadata for audio can sometimes be unhelpful, specifically if you are reliant on a recording of a person being tied to a phone number and a set of metadata. For example, if the phone in a particular office has a number that is delegated to the person who sits in that office, but if that person is out of the office and a colleague decides to make a phone call on that line, the recording made on that phone will be associated with the person who resides in that office. By merely looking at the metadata, there’s no way to determine who made the phone call. In such a case, we can use one of the machine-learning functions to identify nuances
in certain speakers’ voices, pitch and the way that they say things and generate a digital fingerprint of their voice pattern, called diazirination. Using the created voice pattern as a search, you can look across hours of recorded conversations to establish on which calls your target speaker is talking. The product of such a search is a relevance or confidence score from 1 to 100 for all of the recordings, enabling an expert reviewer to listen to a subset of the highly responsive calls and then determine whether the speaker is relevant to the matter or not. Scores can be improved by using speakers from outside the target data set to act as ‘imposters’ to allow irrelevant speakers to be removed from the dataset. Eventually over time, the confidence of the scores increases until the software can confidently identify the speaker’s voice, reducing the risk of mistaken identity in audio review.

Q: How are your clients responding to monitoring requirements?

We see a full array of responses to monitoring requirements. Some businesses see expenditure on compliance as a priority, while others do not. Many businesses are monitoring rigorously and have tool sets and analysts looking at the data, but complacency that regulations are not always clear as to the type of information that they must retain. Few institutions are adequately reviewing their audio evidence. For while there is a duty to make and retain voice recordings, the duty to actively monitor audio evidence is often implemented through employee training and employee awareness as opposed to actual review.

Organizations that have been targeted more frequently tend to take more proactive measures. Over the last two to three years, as a result of the reputational damage that these organizations have suffered, there is more motivation to adopt a more proactive review of their information. However, IT teams see this as a burden. It involves a massive amount of data processing, and organizations don’t always retain the legacy systems to assist with the review and recall of the data. Reviewing every piece of information an organization has can be a colossal exercise. That’s why the tools we have now are making monitoring a reality. Suddenly we are able to search, analyze and review greater volumes of data in a shorter space of time.

Q: What are some best practices for financial institutions?

Financial institutions must have the ability to mine their information and take reasonable steps to ensure their organization is not in breach of the regulations that apply to them. While the care of the data, which includes:

- ensuring that you are not making changes to the source of the data;
- showing that you can explain the provenance of any data;
- attesting to the knowledge of the operators who are applying the process in that they understand the effect that they’re having on the data; and
- more importantly for the regulators and the results, applying the tooling consistently might be customized to the institution’s particular need to connect things internally, all financial institutions should implement industry best practice in the way that they deal with data. The grouping and coding that’s applied to the data must be consistent across the industry so that when workflows and procedures are being documented as they should be, a regulator can be confident that the same standards are being applied across the industry, even if they don’t have the exact same tool sets.

Q: If I transcribe telephone calls, do I have to keep the text for longer than the original?

This is a question that is asked a lot. The advice Epig has been given is that no new business record is created, nor does any specific retention period apply to it, unless you intend for it to replace the original voice recording entirely. In an ideal world, financial institutions would be indexing all telephone calls as they are made, both as text and potentially also using “phonetic” indexing, which would allow greater use of existing monitoring technology, which is very much focused around traditional text media. Epig is well placed to advise on strategies that allow financial institutions to do that in a cost-effective manner.

Q: Why do financial institutions partner with Epig?

The reason financial institutions partner with Epig is because our strengths include identifying the right tools to deploy against large data sets, mining large data sets and producing a subset of valuable data that the compliance team/lawyers need to focus on. Financial institutions don’t want to spend money on the wrong tools and the wrong approach and then overburden their staff resources with manual tasks that an automated process could complete more efficiently and more accurately. Nor do they want to invest in tools that yield a lot of meaningless results, which lead them to think all is well while actually missing items they should be looking at, potentially resulting in regulatory fines and sanctions for failure to adhere to the proper process. Lessons learned in eDiscovery can now be applied to corporate organizations’ monitoring programs. Finding appropriate solutions requires people, processes and technology. A sensible blend of technology and people saves a lot of downstream pain. But technology is not a silver bullet. It’s the tool that allows us to get through huge amounts of data to arrive at a subset of information that requires a more accurate or in-depth investigation. With technological advancements, it is now viable to do that on internal networks.

Q: What’s on the horizon as far as tools go?

Epig is always investigating new tools and approaches. One of the approaches missing at the moment is in the available visual analytics tools. In eDiscovery, it’s in its infancy. In the investigative world, there are more advanced tools capable of analysing unstructured thought. This entails looking at an unstructured thought (e.g., considering whether there is a potential connection between particular groups of people) or an investigative concept and applying it to a massive data universe. The software analyses potential connections and reports that, based on that assumption or concept, the reviewer should also examine groups of related targets engaging in similar conversations. We don’t think this method exists in mainstream analysis at present, but true visualization says that on the basis of what we’ve found, there are probably connections between these people that we haven’t appreciated.

It’s very early days, but tools used in marketing analytics and the introduction of a more cognitive-based analysis and query are moving in this direction. So it is reasonable to predict that that level of cognitive intuition will start to consider where else it has seen something with similar patterns in a data set. We’re already seeing this in market research and in medical diagnosis of symptoms. As that power becomes more readily available and starts to underlie the types of systems that we are using, that will certainly help the investigative process. But the key element is that information is growing at a tremendous rate. Information is very difficult to get into a form in which you can apply that logic, and we are at a point now where the tools are available to get the information stripped, ripped and into text ready to be queried. So we’re starting down that journey with the tools we have today. We don’t advocate waiting for the supercomputer to be developed to do all of this work. Rather, we would suggest that our clients start the process now, knowing that these technologies are coming along and will be able to be plugged into their processes in the future.

Socialize with MCC!