Using artificial intelligence to detect risk in the hybrid workplace: three approaches

Published 12-Apr-2022 by
Stacey English, Theta Lake

The transition to hybrid working has created new and heightened compliance and security risks which are firmly on the radar of financial services firms and regulators alike. The UK's Financial Conduct Authority (FCA) set out one of the clearest mandates that hybrid and remote working environments must have equivalent controls and cannot compromise an organisation's ability to meet regulatory rules or obligations.

Integral to the hybrid workplace are dynamic collaboration tools such as Zoom, Cisco Webex, Slack, and Microsoft Teams. These modern methods for meeting, collaborating, communicating, and sharing information require the same level of oversight as traditional text-based communications like email, which makes the task of detecting risk even more challenging.

As clarification, FINRA has noted that 'visual aids, such as a whiteboard or dynamic charts, or a chat or instant messaging feature during a live, unscripted online conference' may be "correspondence, retail communications or institutional communications, and the firm must supervise them as such."

Similarly, ESMA [page 49] has reminded firms that "electronic communication covers many categories of communications and includes among others video conferencing, fax, email, Bloomberg mail, SMS, business to business devices, chat, instant messaging and mobile device applications."

But it will not produce an exhaustive list given the continuing innovation and advancement in technology - a clear direction that its monitoring and recording requirements extend to all digital communications channels.

However, the video, voice, chat, and file transfer features of these tools introduce myriad opportunities for data exposure such as sharing the wrong screen with personal data showing or attaching sensitive files or links in chat conversations. There are also multiple challenges in detecting misconduct that occurs on screen, or in understanding the true context of communications laden with emojis and images.

The sheer magnitude of communications outstrips the capacity of compliance officers to manually review conversations to spot risks occurring across the business.

The scale and complexity of actively and comprehensively detecting these unique and growing risks has prompted organisations to turn to technology including artificial intelligence (AI) to help. The use of AI enables vast volumes of communications to be analysed. It enables organisations to detect risks and breaches at scale, provides alerts at significant speed and can help prioritise what to review.

Practical compliance implications

When determining the most effective approach to automated risk detection, there are important considerations for organisations to note, ranging from explainability to the level of compliance resources needed for oversight.

Option 1: Word searches

At the most basic level of automated risk detection are word searches, or 'lexicons,' which operate by looking for specific words or sequences of words in communications. These are a long-established and simple approach to detecting specific phrases or words related to a risk. Word searches are highly manual and technically simplistic – they do not incorporate modern AI-based computing techniques, and therefore cannot account for word lookalikes, soundalikes, or the context of conversations.

This basic approach leads to matches being flagged which are not relevant, such as 'I guarantee my son's football team will win' being identified erroneously as a guarantee or promissory statement in the financial services context.

Organisations using word searches for risk detection will either need to dedicate significant compliance resources to reviewing huge volumes of irrelevant 'false positives' or limit their searches to narrower terms in order to reduce the number of alerts triggered. Both options increase the likelihood of missing actual risks.

While this approach is straightforward for compliance and risk professionals to explain, it is difficult to implement in practice because the searches need continual revision as risk-related terms change, such as cryptocurrency developments where terminology is continually evolving. And critically, searches are driven by risks that organisations are aware of, which prevents other risks that are unknown, emerging or occurring within peers or the wider industry being surfaced.

Option 2: General AI detections

The use of AI for risk detection can dramatically reduce the levels of human oversight needed to manually review communications. It provides clear benefits through its ability to understand risks in context, overcoming the limitations of traditional word searches by considering the wider conversation. That said, organisations should pay particular attention to how AI is used for risk detection as all AI is not created equal.

Many AI detection tools are built for general purpose, leaving customers the onerous task of training, and tailoring the detections, which can be a frustrating and time-consuming process. These detections can be challenging to implement because of the considerable time that compliance practitioners need to spend training the data and checking the quality of alerts.

The nature and volume of the datasets used to train the AI also makes it much harder to understand how risk-based decisions have been arrived at, and to explain to regulators, auditors, boards, or other stakeholders. While these detections will ultimately enable an organisation to pinpoint the risks it knows or is concerned about, they create a significant blind spot because new, emerging, or unknown risks will not be surfaced for attention.

Option 3: Industry specific detections

By contrast, organisations have the option to deploy purpose-built, pre-trained AI-based risk detections which focus on specific conduct, compliance, or security risks. These purpose-built detections use relevant, high-quality datasets, which allows for accurate detection of complex, industry specific notions like insider trading, collusion, or suitability.

The models can be trained to detect specific confidential or personal information including account numbers, email addresses, and birthdates, sensitive documents like customer lists or applications that are shown such as trading screens, HR or finance systems.

These targeted detections use high quality expert sources and domain expertise, which means that the burden does not fall to individual organisations to train the AI models or verify the results. The focus on specific data sets and risks is also more transparent making it easier for practitioners to explain how risks are identified and what triggers an alert.

Its ability to understand specific risks in context reduces both the number of false positives or alerts that are not relevant, as well as risks that would otherwise be missed because audio or transcript are unclear.

For example, if a conversation was transcribed and included the words 'count member dirty won for you too,' by analysing the proximity and context of other words, this AI-enabled approach would disambiguate the dialogue and identify it as a conversation about 'account number thirty-one forty-two.' A simple keyword search of a transcript would not accurately identify it.

A significant benefit of this approach is the small number of false positives or irrelevant alerts, which ensures that valuable compliance resources can concentrate on addressing actual risks occurring within their business. Critically, this approach ensures that risks emerging across the industry or other similar organisations will be flagged, which would otherwise be missed in custom built detections trained by individual organisations.

Making the right choice

With growing instances of data loss, cyber-attacks, and misconduct in the hybrid workplace, choosing the most effective approach for comprehensive risk detection is paramount to protect the business and its customers. In an environment where it has never been so easy for anyone to digitally record and share externally what is happening across modern communications, it is essential for organisations to be on the front foot in detecting and dealing with potentially reputationally damaging risks, ranging from data loss to threatening behaviour.

Comparable to finding a needle in a haystack, AI enables organisations to find the risks across its communications at speed, and benefit from significant efficiencies and cost savings.

But in choosing the most effective approach to risk detection, organisations should take care to ensure their valuable compliance resources are not diverted to training and checking AI models, and can instead focus on reviewing and dealing with any important risks and issues that are flagged.

Stacey English *is director of market intelligence at Theta Lake*

Complaints Procedure

Produced by Thomson Reuters Accelus Regulatory Intelligence 12-Apr-2022