A 1LoD publication

in-Focus®

REPORT

Comprehensive Surveillance and Risk Detection: Multilayered AI Coverage

Commissioned by:

VERITAS™

“With AI/ML and Classification we reduced the number of alerts over 80% while simultaneously increasing the number of findings requiring remediation by 15 times. This high number of violations led to firm-wide re-education programmes and a subsequent lowering of risk across the firm.”
- Global Financial Client Using Veritas

Modernize workflows, reduce noise, and increase findings.

The hype around ChatGPT is maybe overdone, but AI and large language models will certainly transform thousands of processes across almost every industrial sector. In financial services, one obvious area of impact is in communications surveillance but the devil is in the details: what type of AI model should you use and why? What is TAR 2.0 and what makes it different to other forms of AI? And is it a solution to the problem of noise in communications surveillance?

The challenges faced by communications surveillance teams are all too familiar. Simply capturing all the channels on which regulated business is occurring has proved problematic. Recent enforcements demonstrate the need to actively capture all forms of communication used by broker/dealers and this extends to being aware of potential ‘off-channel’ communications and taking corrective action.

As electronic trading venues, video collaboration tools and messaging services proliferate, that capture challenge will just keep getting bigger. Turning the huge volumes of data ingested from these channels into useful raw material for surveillance engines taxes even the best-resourced organisations, and the output of these engines is a firehose of alerts, alerts that are overwhelmingly just noise.

Is new technology the answer? Surveillance chiefs are wearily familiar with the claims made for it and point out that while the theoretical benefits are large, achieving those benefits in the hugely complex environment of a large financial institution is much more difficult.

However, the advent of large language models and other techniques in artificial intelligence is a gamechanger, particularly in the area of communications analysis and alert accuracy.

The larger, best-practice banks in the industry have already adopted a number of different types of AI at various points in their non-financial risk management processes, and it is becoming clear that AI has quickly gone from ‘nice to have’ to ‘must have’.

Why models matter

Proliferating communications channels require smart automation from ingestion to review. Models designed by data scientists are one answer, but they require significant initial training and suffer from model drift. Another approach is to use the review process itself as an input into machine learning models. This approach, known as Technology Assisted Review (TAR), is an iterative process where a subject matter expert (SME) reviews text samples and then the model applies coding to the total dataset based on what it has learned from the samples. That process is repeated until the model is able to identify pieces of text relevant to the review process to accuracy levels acceptable to the user.

In its original form, sometimes called TAR 1.0, this is a multi-step process involving between six and 10 steps. These include reviewing randomly selected samples multiple times. Because the samples are random, they include low-value items – in surveillance this would mean samples that lead to false positives or to no alert at all – which has two negative effects. First, it increases the number of iterations (and so time) required to achieve the desired accuracy. Second, because the samples are randomly selected, human reviewers do not have the opportunity to provide feedback to the machine learning models by introducing documents of a higher value that are outside of the prescribed samples.

TAR 2.0 – sometimes known as continuous active learning – solves these problems. Here there is no multi-step process. Analysts simply point to items that have already been reviewed. These items have already been classified by the team and so the training set for the model is that reviewer history. There is no need to build and cull a special training set or tune the engine. The model then updates itself regularly, every week or every day or even more frequently.

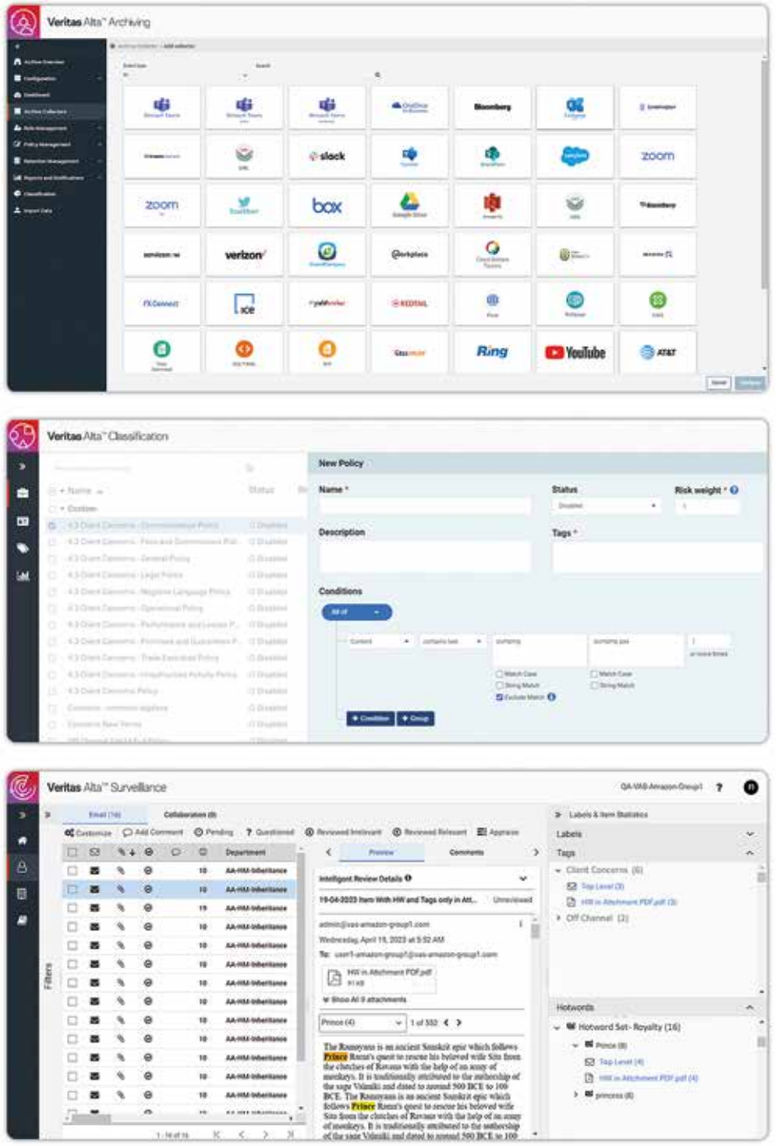
As one proponent of TAR 2.0 explains, “As information comes in and is reviewed, we apply machine learning – but not the usual ML used by others, we took a technique from the e-discovery realm of continuous active learning as opposed to a data scientist-led model building. We chose to go down the TAR 2.0 approach or so called econtinuous active learning as it’s called an e-discovery world, where the expert is the reviewer on the front line and that is the training.”

The Veritas approach is different in that it learns from what the reviewers themselves are doing so if there is a change, then that will be reflected in the human reactions and instantly incorporated into the models.

There are pros and cons but the first type of model requires billions of iterations and a huge dataset to train it and if new information enters the system then you can get model drift. “Ours is different in that it learns from what the reviewers themselves are doing so if there is a change, then that will be reflected in the human reactions and instantly incorporated into the models. We also allow reviewers to submit material to the model for it to learn from – for example, words or phrases that they have found are associated with dubious behaviour – so reviewers can essentially enhance lexicons or the alert engine ML in real time.”

In essence, the machine simply observes the analyst team in action and learns from its members directly and continuously. There is no need for a subject matter expert or a data scientist to tune the engine. It’s all built right into the system. This means that the system never goes stale – the model never drifts. In fact, it gets more accurate over time. The more data you feed it, the more accurate it becomes.

For surveillance and supervision, as there are always new data, new emails, new challenges and violations, TAR 2.0 is a significant development.



Explainability built in

A significant benefit of this model – beyond the absence of model drift – is that this process delivers explainability. It is always clear how the model has reached its decision, and any particular alert score can be interrogated for the relevant and irrelevant items that generated it.

What about ChatGPT?

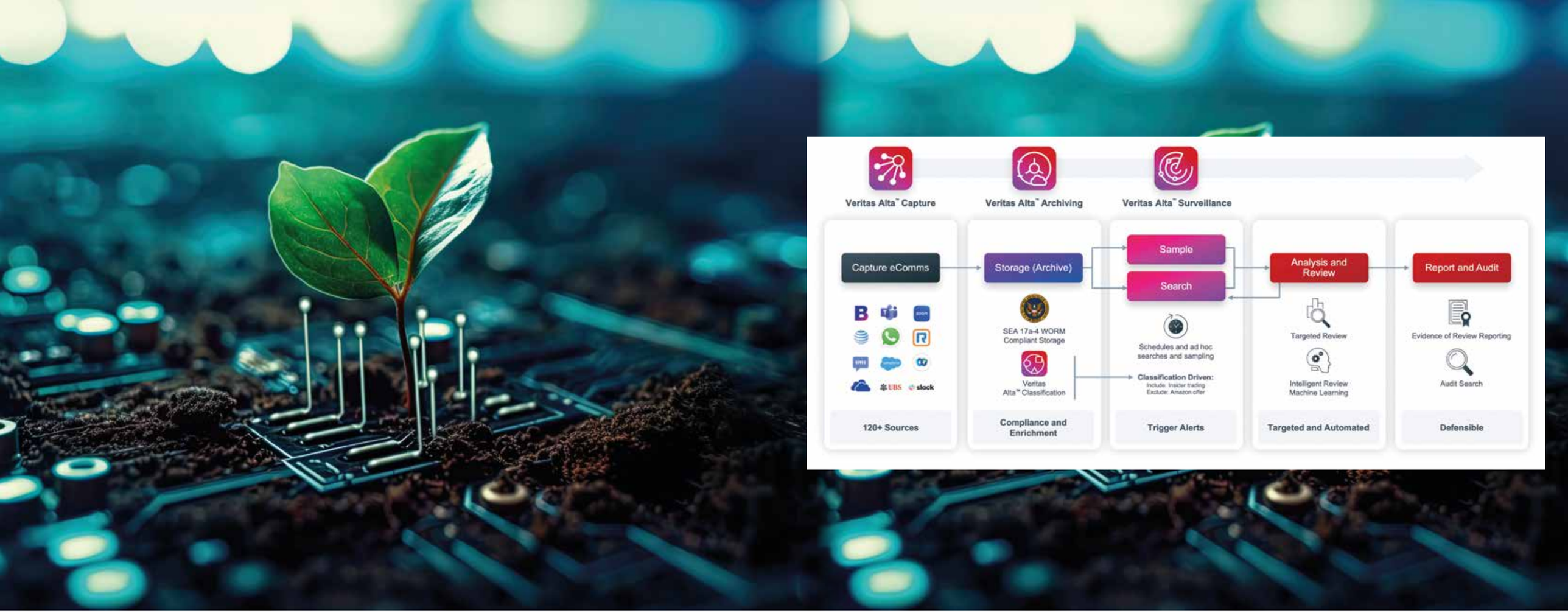
Another way to analyse text and identify items relevant to search criteria is to use a large language model (LLM) such as ChatGPT. These work differently than TAR because they use natural language processing to learn language patterns and structures and the relationships between words. But like TAR these models do not possess general intelligence. Their abilities are limited to the patterns and information they have learned during training.

Used carefully, these types of models can deliver impressive results, but if used by themselves they can create problems in an environment such as communications surveillance. Augmented AI is the answer.

First, relying completely on this type of AI creates a single point of failure: if the model gets something wrong, there is no additional corrective mechanism or feedback loop to mitigate the error.

Second, it is difficult to prevent AI models trained on vast datasets to reflect the biases in those datasets. Those biases will distort the models’ outputs, and in a use-case such as communications surveillance those distortions are likely either to cause increased false positives or, worse, to miss true positives that the model cannot identify.

And third, AI-only models tend to be less explainable than hybrid ML/assisted review models. This makes them less attractive in sectors where regulators demand explainability. It also makes them harder to review for bias: if a human reviewer can understand how the model arrives at its decisions then they can more easily identify biased reasoning.



The right solution for surveillance?

How are these technologies of practical use to communications surveillance professionals today? They need to capture and store all of the data required by regulators – including any channel used by any regulated person to conduct regulated business.

One example is Veritas Alta™ Capture which can ingest 120 different channels including email, messaging and other e-comms. Partnerships in voice give it the ability to ingest and transcribe audio and expand sources to include even obscure content sources not supported by other vendors.

But capture is just the start. Banks also need tools to help them manage the growing volume of data and the number of (mostly trivial) alerts that analysts face.

To manage today's data volumes and to reduce false positives, Veritas applies automation and AI across the whole ingestion and surveillance process to improve efficiency and effectiveness. It uses machine learning in two different ways at different stages in the process.

The system incorporates a 'pre-trained' element consisting of an automated classification engine based on 200 pre-trained expert-reviewed policies leveraging NLP, keywords and nested condition logic that assign data to different classes.

It uses this engine in the indexing and classification stage of data ingestion. Veritas enriches data and metadata as it goes into the archive, and the ML and NLP in the classification engine can also carry out functions such as language detection and sentiment analysis.

Then in the monitoring/surveillance process itself, the system additionally uses the TAR 2.0 methodology to reinforce the model with the external information generated by the human alert review process.

Finally, the system can use its embedded ML to help with compliance reviews – to ensure that the surveillance process is aligned with enterprise and regulatory policy. And it embeds reporting and audit tools that provide assurance to the 3rd line that the firm has a process in place to supervise e-comms and can report on its activities.

For more details on the capabilities of Veritas Alta™ Surveillance, read this data sheet.

If you are interested in further discussion, please contact Veritas today.

Efficiency, effectiveness and explainability

Banks can no longer afford not to have a capture and archiving solution capable of ingesting the many dozens of feeds and media that are now irrevocably part of their trading, wealth management, retail and corporate banking activities. They also need a solution to the inefficiencies, costs and analyst fatigue caused by false positives in surveillance itself.

Many banks would prefer a holistic, integrated solution rather than a hard-to-manage and complex stack of solutions each of which addresses a particular medium or market type. But now they also want as much automation as possible in the initial classification of data and in surveillance models and alerting to reduce costs and increase efficiency.

These solutions exist today. And by combining cutting-edge machine learning and artificial intelligence algorithms with the data from human review processes, they deliver efficiency, effectiveness and explainability.

1LoD®

www.1lod.com