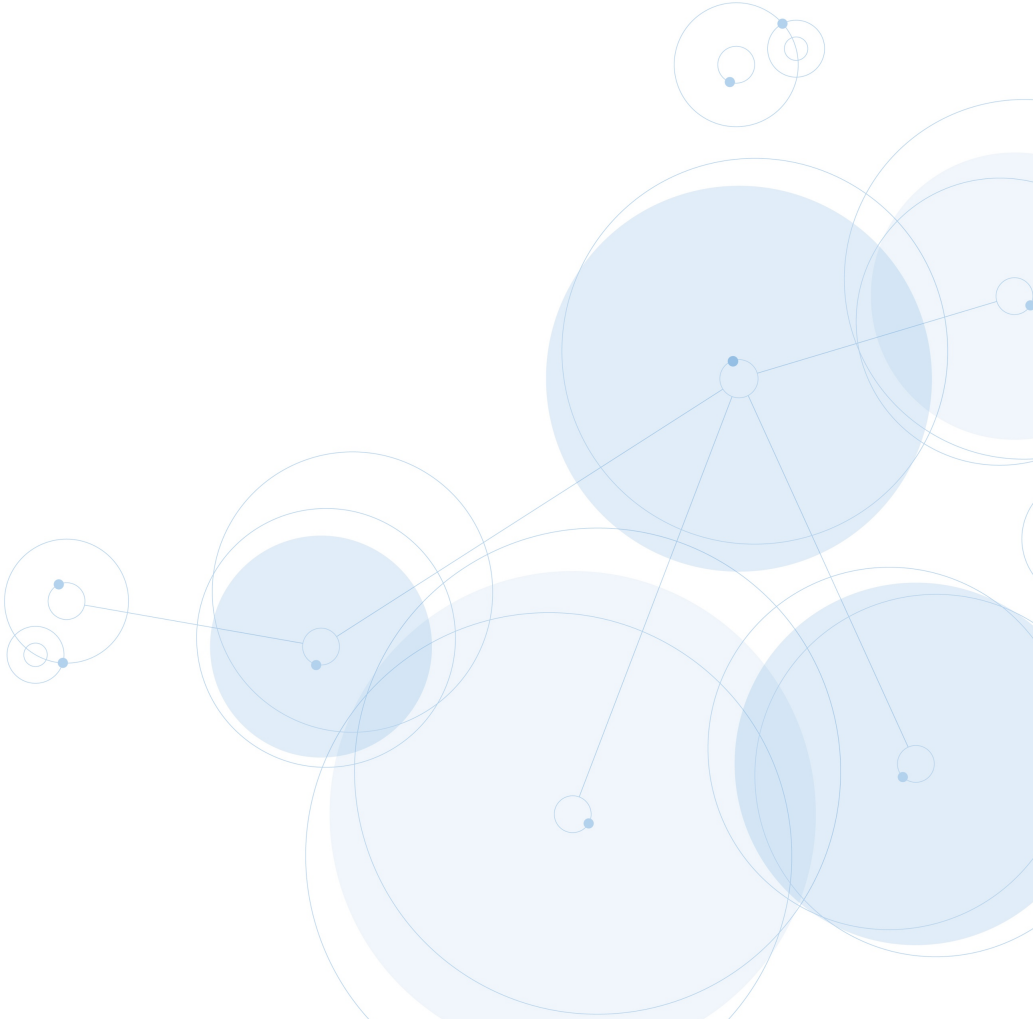


AYASDI

Anti-Money Laundering Solution Deep Dive

WHITE PAPER



An AI-Driven Approach to AML

Anti-Money Laundering (AML) is a particularly challenging area of regulation for banks – even more so for large, geographically diverse institutions. Failure to have adequate processes can result in massive regulatory fines.

Preventing money laundering in the financial systems has high stakes. It aids in the prevention of terrorism, human trafficking, and narcotics distribution. What makes matters worse is that existing AML systems and processes have proven to be drastically inefficient. Proven by the fact that only 2-5% of global GDP, roughly \$1-2 trillion annually, is made of money laundering transactions.

Ayasdi's solution for AML delivers order of magnitude improvements by using artificial intelligence (AI) in ways that dramatically enhance existing technologies and processes. The precise application of these technologies delivers extraordinary results without the need to engage in large scale software upgrades. As a result, leading financial institutions around the globe have deployed Ayasdi's AML application against their most complex challenges.

Finding the Leverage

What makes AML such a difficult problem to solve is that it involves complex data, detailed workflows, and significant human involvement. The result is that the cost of compliance is increasing by 50% year-over-year and quickly becoming a drag on earnings at a critical time for financial institutions. The dimensionality of the data and the significant costs to support a robust AML process are amplified for large, geographically diverse financial institutions.

A current AML process generally looks like the following:

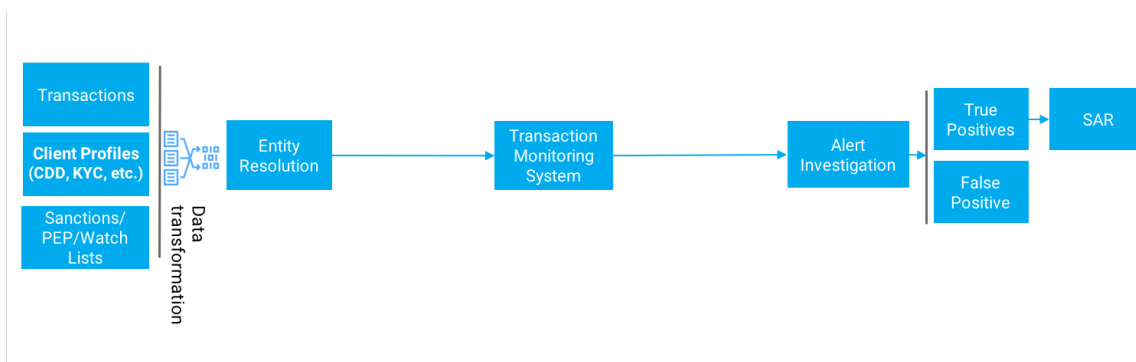


Figure 1: A typical AML process inside of a financial institution.

The major issues are the volume of investigations, the prioritization of those investigations, and the classification of the typologies that feed those investigations. These issues are not effectively addressed in most AML processes – even sophisticated ones.

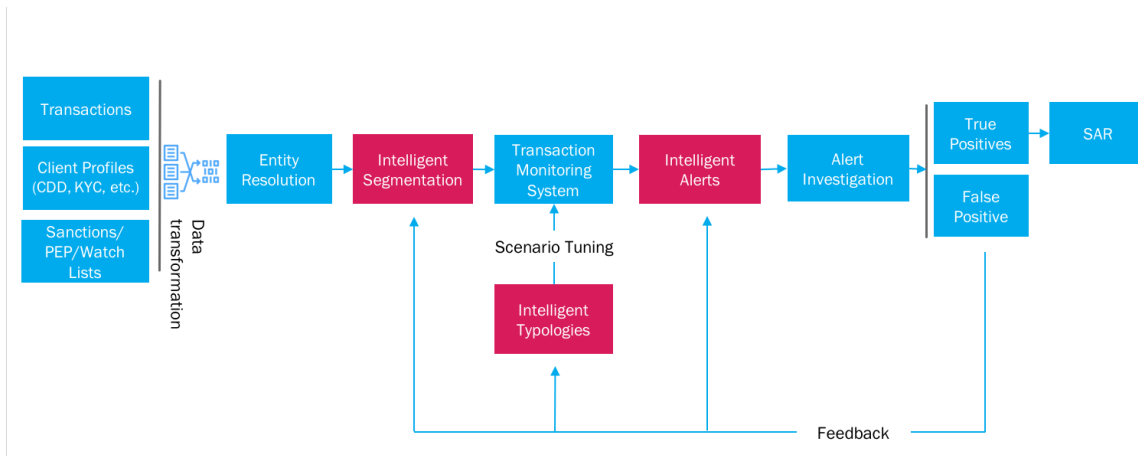


Figure 2: An augmented AML process inside of a financial institution applies intelligence at key lever points to produce orders of magnitude better performance.

The challenge begins with the volume of investigations and the fact that as many as 95% of these investigations do not result in a Suspicious Activity Report (SAR). This means that 95% of the effort of a large investigations team – which can contain anywhere from 500 to more than 5,000 people is not required.

At the heart of the problem is finding the balance between signal and noise. Too much noise in the form of false positives increases costs for the bank as they need to investigate the suspicious activity reports. Too little signal means that the bank is exposed to regulators because they might miss criminal activity.

Traditional AML methods have not scaled to the current regulatory environment. For example, most AML processes typically have hand-coded money laundering rule patterns, known as typologies to evaluate each transaction for each geography or type of business. Further, most AML triage programs use LIFO or transaction amount to prioritize investigation. Finally, typologies, are coarse at best, and non-existent at worst.

Financial institutions need to re-imagine these processes using artificial intelligence, and such reimagining requires a principled approach.

A System of Intelligence

From financial crimes to regulatory risk, intelligent applications offer the optimal blend of power and breadth. What, however, constitutes an intelligent application? For an application to be intelligent it needs to incorporate the following five capabilities:

1. Discovery
2. Prediction
3. Justification
4. Action
5. Learning

Let us consider each of these concepts in turn.

Discovery

Discovery is the ability of an intelligent system to learn from data to support the development of applications. Often this needs to be done without being presented with an explicit target. It relies on the use of unsupervised and semi-supervised machine learning techniques (such as segmentation, dimensionality reduction, anomaly detection, etc.), as well as supervised techniques where there is an outcome or there are several outcomes of interest.

In complex datasets, it is nearly impossible to ask the “right” questions. To discover the value that lies within the data one must understand all the relationships that are inherent and important in the data. That requires a principled approach to hypothesis generation.

Ayasdi’s unique technology, topological data analysis (TDA), is exceptional both 1) at surfacing hidden relationships that exist in large and complex data sets, and 2) identifying those relationships that are meaningful without having to ask specific questions of the data. Ayasdi’s output represents complex phenomena visually, and therefore is able to surface both weak and strong patterns or relationships in the data. This permits the detection of emerging phenomena that may otherwise remain hidden.

As a result, enterprises can now discover answers to questions they didn’t even know to ask.

Prediction

Once the data set is understood through intelligent discovery, supervised approaches are applied to predict what will happen in the future. These types of problems include classification, regression, and ranking.

To implement prediction capability, Ayasdi uses a standard set of supervised machine learning algorithms including random forests, gradient boosting, and linear/sparse learners. The discovery capabilities of Ayasdi's technology are highly useful in that they generate relevant features for use in prediction tasks or finding local patches of data where supervised algorithms may struggle.

Justification

No manager responsible for AML operations should deploy intelligent and autonomous applications against their AML process flow without a thorough understanding of what variables power those applications.

An AML process that employs AI needs to reveal its workings in a way which that makes outcomes recognizable and believable to its entire constituency. For example, when one recommends a segmentation of customers, it is important to have an explanation of how the machine is making that determination in ways that are interpretable by humans. This kind familiarity is important in generating trust and is particularly relevant in situations where a regulator is involved.

Action

The process of operationalizing an intelligent application within the enterprise requires some change in the organization, an acceptance that the application will evolve over time and that will demand downstream changes – automated or otherwise.

For this to happen, intelligent applications need to be “live” in the business process, seeing new data and automatically executing the loop of Discover, Predict, Justify on a frequency that makes sense for that business process. For some processes, that frequency may be quarterly; for others, daily. That loop can even be measured in seconds.

Intelligent applications are by their very nature designed to detect and react when data distributions evolve. These applications need to be “on the wire” in order to detect evolving phenomena before they become problematic.

Too many solutions provide an answer at a single point of time. A truly intelligent system is one that constantly learns by utilizing the framework outlined here.

Learning

Intelligent systems are designed to detect, react, and improve as the input data stream evolves. An intelligent system is one that 1) is always learning, 2) operates live within the workflow, and 3) is constantly improving. In the modern data world, there is no room for an application that is incapable of learning.

By applying this intelligence framework to the areas outlined above (segmentation, alerts and typologies) an institution can make massive gains in efficiency and risk mitigation. Let us take them in turn.

Intelligent Segmentation

The false positive problem in AML is primarily a function of poor segmentation of the input data. Even sophisticated financial institutions using advanced machine learning techniques for segmentation can suffer from non-uniform groups that generate false positive rates above 95% as well as false negative rates that exceed what most regulators would be comfortable with.

Traditional segmentation is typically static and coarse since it only takes into account a limited set of factors. Moreover, segmentation is done manually and separately for customer data and transaction data. The result is an inability to effectively capture complex feature interactions.

These weaknesses in segmentation have a massive effect on downstream operations and processes when coarse segments and inflated thresholds hit the transaction monitoring system.

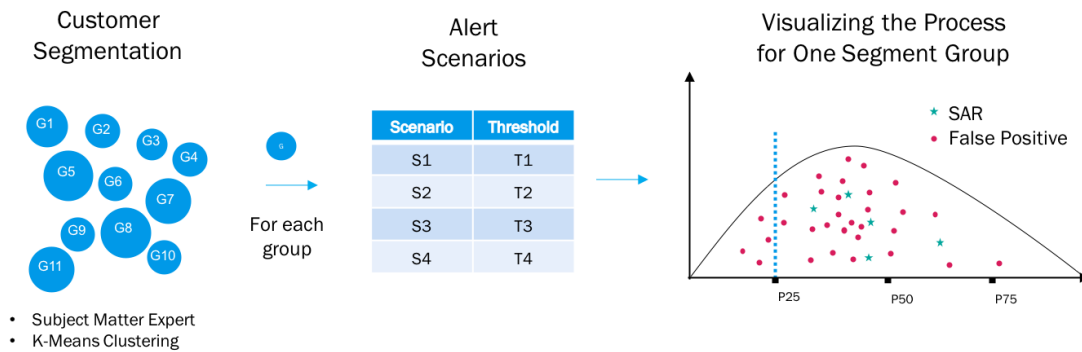
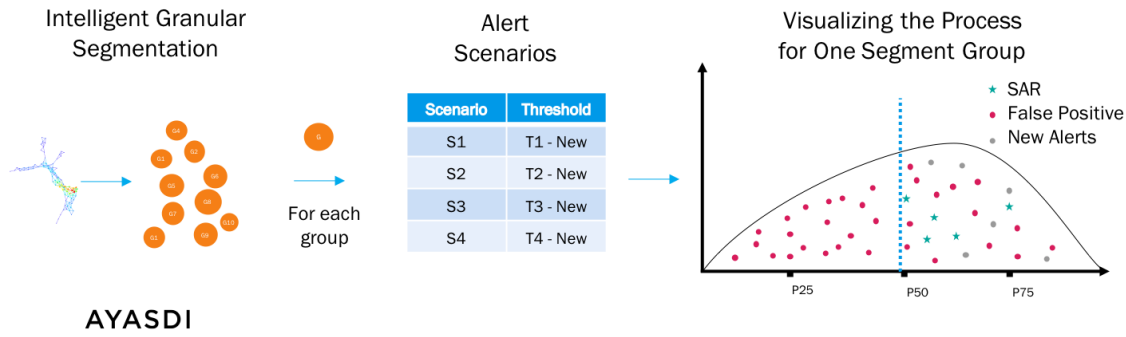


Figure 3: A typical segmentation process produces uneven groups and this means that thresholds must be set artificially low – resulting in a significant number of false positives (shown in red).

When AML rules are triggered, the offending transactions must be investigated and adjudicated. The result can be a Hobbesian choice of either regulatory exposure or throwing bodies at the problem.

Ayasdi has developed an approach that automatically assembles self-similar groups of customers and customers-of-customers. This results in more granular and uniform segments and thus in higher thresholds.



AYASDI

Figure 4: An intelligent segmentation process delivers far more granular and uniform groups, resulting in higher thresholds and fewer false positives. In addition, these granular groups catch the false negatives too.

Further, Ayasdi’s approach has an insatiable appetite for data. The more data sources that are available, the better the grouping that results from Ayasdi’s approach to segmentation. The sources in this scenario include both customer and transaction data. More importantly, Ayasdi’s technology does not require labeled data in order to derive the initial segmentation. Removing the requirement for labeled data permits substantial expansion of the number of data sources and means that the customers of a bank’s customers (KYCC) can be considered.

Equally important is the fact that the Ayasdi approach provides complete transparency into what drives the segmentation. Further, Ayasdi’s platform produces a complete documentation workflow containing simple decision trees that can be shared with internal model governance boards and with external regulators.

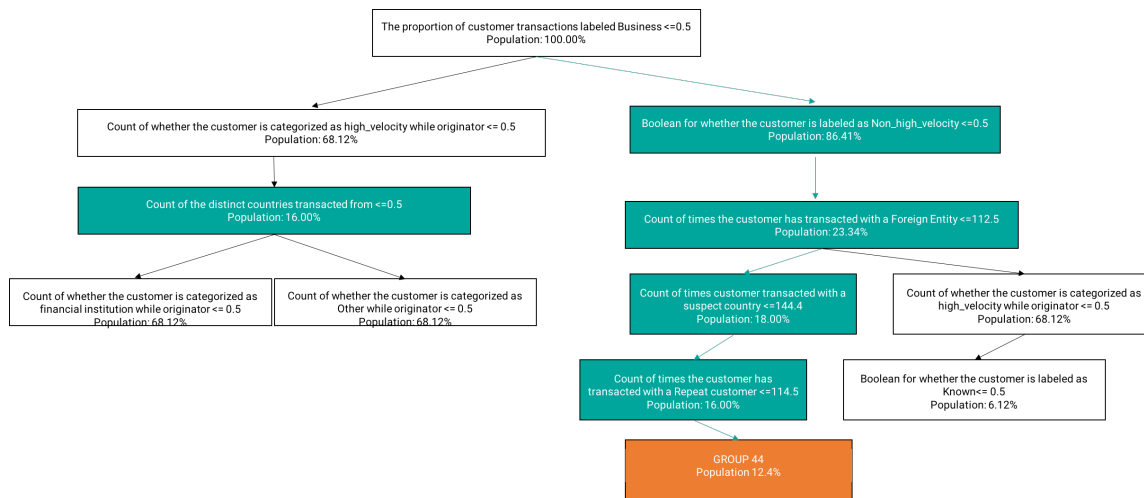


Figure 5: Decision trees are excellent ways to visualize complexity for regulators and internal model review boards and are a key part of the justification step in AML.

The solution, however, is not static. Rather, Ayasdi’s AML solution constantly looks at newly arriving data, identifies changing patterns, and suggests updates to segments and rankings based

on that new information. As a result, subtle patterns suggesting emergent behavior are readily identified for consideration by subject matter experts.

Intelligent Alerts

The term “Intelligent Alerts” refers to the ability to accelerate clearing of an alert backlog by automatically categorizing alert priority (L1, L2, etc.). A critical part of this capability is providing the reasoning for an alert’s auto-dispositioning. Alert auto-dispositioning is important because it allows investigators to focus their attention on the highest probability items while relegating the lowest probability items to the bottom of the priority list.

Below is a description of how the Ayasdi approach to Intelligent Alerts works.

Like most AI problems, Intelligent Alerts use two “patterns,” hotspot detection followed by ranking. This process starts with the output of the transaction monitoring system, which is a set of alerts (A_1 through A_N). These alerts are first grouped using Ayasdi’s Topological Data Analysis technology based on the notion of similarity. The result is a set of Groups (G_1 through G_N).

With the alerts placed into Groups (G_1 through G_N) we then move to ranking.

Within each Group the alerts are then assigned a level using a methodology that grades suspiciousness. Most financial institutions use three such lots, L1, L2, L3, so we will follow this approach, although other methods work as well. For those unfamiliar with the concept, L1 alerts are closed (not suspicious) with little effort, L2 alerts are closed with some effort, and L3 alerts are filed as suspicious. The output of this operation is Groups containing alerts that are ranked based on the probability that they will result in a suspicious filing (L3 on top, followed by L2, then L1).

While the initial groups are established using the concept of similarity, the alert assignments (to L3, L2, L1) require back testing with historic data. For a typical bank this task utilizes a month or so of data, however it is statistical in nature and the amount of data used will vary.

One additional benefit becomes apparent during the Suspicious Activity Report filing process. By grouping like Alerts together, astute investigators can bundle multiple grouped and linked alerts into one SAR, thereby providing more context for enforcement agencies and aiding them in better catching money launderers.

By using this approach, a bank learns where to focus its efforts. This results in a lower risk profile and increased efficiency.

Intelligent Typologies

Intelligent Typologies, like Intelligent Segmentation, is another way to get the most out of the Transaction Monitoring System while avoiding costly, lengthy or complex system changes. Both create an intelligent, flexible framework from a static, rule-based system. The Ayasdi implementation of Intelligent Topologies works as described below.

Intelligent Typologies are based on the concept that groups of like people and like companies act in predictable ways. For example, the financial transactions for one pharmacy may look similar to those of other pharmacies. However closer inspection may reveal suspicious characteristics like having many larger than average transactions and no health insurance payments. These characteristics are, by definition, anomalous and worthy of investigation. The challenge, however, is that an investigator must search for these exact characteristics among trillions of potential combinations.

Intelligent Typologies draw on two patterns that dramatically improve upon the human-driven approach to inquiry. The two patterns are anomaly detection and ranking. Like segmentation, they are applied in sequence.

Ayasdi's Topological Data Analysis creates networks based on the concept of similarity. It is also exceptional at detecting the opposite of similarities, i.e. anomalies. TDA does not need to be trained using data with a known outcome. TDA doesn't even need labeled data to identify those entities that qualify as anomalous. Furthermore, the Intelligent Typology approach learns over time, demonstrating far more flexibility than do traditional rule-based systems.

Once those anomalies are detected, an Intelligent Typology approach then ranks the anomalies - from least likely to occur to most likely to occur. Banks typically target the least likely or most anomalous first.

Only through Intelligent Typologies can financial institutions catch the ever-changing behavior of money launderers. Furthermore, Ayasdi's Machine Intelligence Platform is able to go one step further by providing full transparency to the process by providing the reasoning behind the anomalous behavior. This transparency is critical to winning over regulators and law enforcement.

Case Study - AML Detection for Correspondent Banking

Doing AML well is challenging enough when the customers are your own. But what about your customers' customers (KYCC)?

Ayasdi was asked to tackle AML for one of the world's largest, most geographically diverse financial institutions. The bank had a goal to improve operational efficiency in the KYCC area by 3% with a stretch goal of 5%. Using Ayasdi's AML solution, the bank achieved reductions in investigative volume of more than 20% while lowering their regulatory exposure by discovering new risk segments that had previously gone unnoticed.

Here's how Ayasdi did it:

We first looked at the available Swift message data. Ayasdi had access to the same features as the client, however, our technology was able to use more of these features and do so more intelligently. The increased feature count was a function of being able to incorporate and create features such as transactional data (type, direction, value), customer data (geographical, chronological), and risk data.

Ayasdi then created a series of prospective segments using a subset of the data. Notably, Ayasdi did not increase the group count, which was kept constant. Rather, Ayasdi created more intelligent, defensible, and uniform groups that were constructed using completely different features from the bank-constructed groups.

It is particularly notable that this exercise was done entirely without supervision. Ayasdi's software selected the appropriate algorithms, created candidate groups, and tuned the scenario thresholds within those groups until the optimal ones were identified.

Once the optimal group structure was identified, a decision tree model was created. The distribution of customers within these groups was then evaluated, validated independently, and deployed against the bank's existing infrastructure.

This entire project took a matter of weeks and involved a team of two data scientists and a project manager from Ayasdi working with a team of two domain experts from the bank. Following a review by the bank's internal model review board and the regulator, the bank is in the process of deploying Ayasdi's segmentation solution globally.

Summary

In the evolving world of regulation there is considerable pressure to achieve a target of “zero failure.” The approaches that banks have employed to date, both historically and in response to enhanced regulatory scrutiny is fundamentally driven by human power. The complexity of the AML challenge, however, is not well suited to the hand-coded approach that dominates most institutional thinking.

Ayasdi’s machine intelligence technology represents an opportunity to apply sophisticated AI techniques to this mission-critical regulatory function. Application of machine intelligence to AML processes dramatically improves efficiency while simultaneously reducing regulatory exposure. Furthermore, this new generation of technology provides exceptional transparency with respect to the creation of rule sets and the process of segmentation.

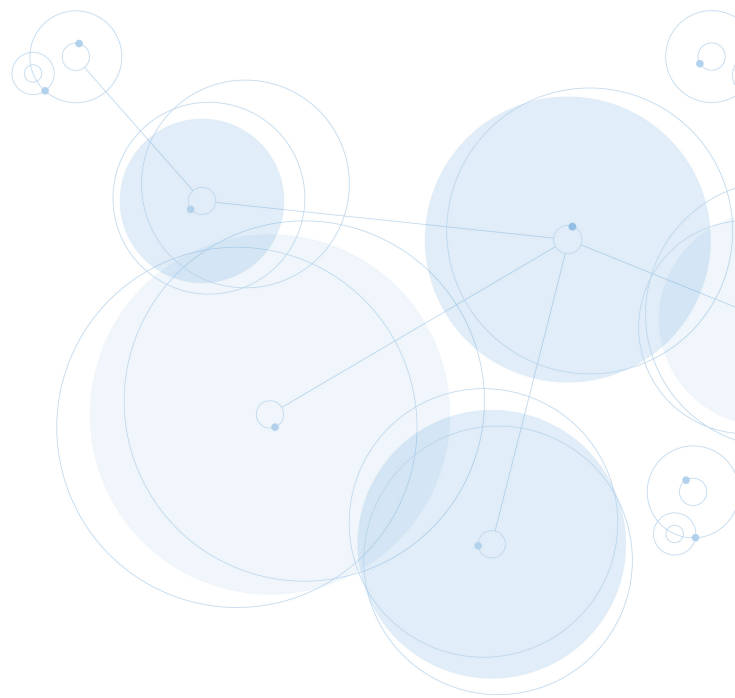
Ayasdi helps propel institutions toward this goal now.

To find out how to leverage AI for your AML challenges contact us at sales@ayasdi.com to arrange a demonstration.

About Ayasdi

Ayasdi is the global leader in the development of enterprise-grade, machine intelligent applications for financial services, healthcare and the public sector. Powered by breakthroughs in both mathematics and computer science, the company's software platform has already solved some of the world's most complex challenges.

Based in Menlo Park, CA, Ayasdi is backed by Kleiner Perkins Caufield & Byers, IVP, Khosla Ventures, Centerview Capital Technology, Draper Nexus, Citi Ventures, GE Ventures and Floodgate Capital.



AYASDI

4400 Bohannon Drive, Suite 200
Menlo Park, CA 94025 USA
+1 650.704.3395
sales@ayasdi.com
ayasdi.com | @ayasdi