Data quality, AI and the future of financial crime compliance

Why banks and credit unions need trusted data, proactive screening and defensible AI

Ahead of the Curve, a Bankers Podcast

This transcript has been edited for clarity and readability.

Kate Randazzo:

This is Ahead of the Curve, a Bankers Podcast. On today's episode of Ahead of the Curve, we're talking about sanctions compliance. We'll discuss why expectations around data quality and financial crime compliance are especially critical right now as geopolitical volatility increases and OFAC updates become more frequent and complex. Regulators expect stronger, more defensible screening programs. For savvy banks and credit unions, that means using AI-powered screening tools to do more with leaner teams.

I'm joined by Greg Pinn, who has spent more than 20 years building software and data solutions to solve AML and financial crime challenges. At Abrigo, he leads the development of innovative solutions such as sanctions and watchlist screening offerings, crypto compliance and risk data products that help banks and credit unions navigate these operational demands.

We are also joined by special guest Sarabjeet Singh, founder and CEO of Resolute. Sarabjeet has designed and operated large compliance centers of excellence for global banks and fintechs, including multiyear programs for top Wall Street banks and global payment providers, combining operational scale, machine learning and automation. His teams have built and maintained global AML datasets for more than a decade, and we're thrilled to have his expertise today.

Sarabjeet Singh:

Absolutely. This is very topical, and we see more and more clients and users looking for more updated data on an ongoing basis. What was good enough from a recency perspective some time ago is not good enough now. The need for data to be defensible and complete is clear.

The need is not just breadth of information, but depth. That includes ownership layers. There is a need to go deeper, so the quality of data around those layers starts to matter. What is publicly available, and can you bring that information to the fore?

Lists also need to be more current and continuously updated from a data-intelligence point of view. That clearly affects data quality. In fact, many tests clients conduct when evaluating Resolute data focus on recency and on subjects that may have been missed on less common lists, especially when those lists are topical in a particular jurisdiction. They are also looking at the coverage.

What we see is that it is not just direct exposure, but indirect exposure as well. The connections and relationships that data can reveal start to matter quite a bit. That is how I see the change

happening. This has been true over the last several years, but especially now, when everything is connected, there is so much data available, and there are so many tools available to sift through it. That's how I look at it from an expectations point of view.

Greg Pinn:

Yes, and when we look at this from the community bank and credit union market, we're seeing a lot of global data, global lists and a huge world of PEPs. The big struggle that many community banks and credit unions have is sifting through that information and making sense of it.

Citibank, JPMorgan Chase and Barclays, these huge global banks, have armies of compliance professionals. They may have a data team that figures out exactly what they want to screen, along with an army of people who can screen millions of PEPs, thousands of lists and hundreds of sanctions lists.

For the community banking industry, this is much more about identifying where the risk lies for that institution. The pressure will continue to move from larger banks down to community banks, credit unions and local banks. They need to do more screening, understand who they are doing business with, and take the know-your-customer concept much more seriously.

A lot of that means understanding, at a high level, who your customer is, figuring out the risk profile of your customer cohort, and then determining the correct data to screen for that cohort. How do you balance risk and resources? In the 20 or 25 years I've been doing this, that tension has been part of every institution's work. Everyone wants to be more risk-averse, and everyone has limited resources. That's true at the high end with global banks, and it's true with local community banks.

The big difference is that local banks need a more laser-focused, strategic approach, whereas larger banks can be much more diverse. That's a lot of what we're seeing from a data perspective: a shift away from, "Hey, this is just all about OFAC," and toward, "I need to understand my risk."

Kate Randazzo:

That makes sense, and I like what you both said about pressure increasing over the past few years. In the past, we've seen small banks think, "I'll do this for regulatory purposes, but it doesn't apply to me as much as it does to large banks."

If having a unified, high-quality data set across PEPs, sanctions and watchlists is critical today, but smaller institutions do not have equal strength across all three, what would you say to them about why that is more important now than ever?

Greg Pinn:

I think this is all about being proactive. On the transaction-monitoring side for Abrigo customers, that may be BAM+. In fraud detection, in our world, it's AFD. Regardless of the tools you use, we see a huge focus on transaction monitoring and fraud detection.

Those are both reactive. They respond to events that have already happened at your institution, such as a customer involved in kiting or structuring, an account takeover or another type of fraud. Proper list screening is proactive. It identifies the individuals coming into your institution and the current or potential customers who pose a risk to your institution.

Consider a bank that specializes in a specific industry, geography or community. One of the great things about the U.S. banking system is the sheer number of banks and the history of banks that support immigrant communities. We have banks that support Chinese, Filipino, Mexican, Central American, Indian and Korean immigrant communities, as well as communities from around the world.

This helps people who are new to the country get access to financing, loans and the banking system. It is a key part of welcoming people into the United States. The challenge for a bank is understanding the risk of people who do not have a history in the country.

If I do a lot of business with the Korean immigrant community, I need to make sure I'm screening lists that are local to Korea. I need to understand whether the origin of the people coming into my institution creates risk outside the United States and outside the context of sanctions. I want to get ahead of any associated risk.

That is not just true for immigrant communities. It is also true for banks that target specific industries. Some financial institutions focus specifically on the medical community, helping doctors, physicians and hospitals get loans. Those institutions need to understand who has been involved in Medicare, Medicaid and insurance fraud.

Other financial institutions focus on the energy sector. There are lists associated with environmental crime, including information from agencies such as the EPA and NOAA. This is the type of work institutions need to do to get ahead of risk and reduce the amount of work they do on the reactive side.

You're never going to get rid of all fraud or all high-risk transactions, but you can proactively identify the high-risk people within your community and identify high-risk transactions before they become a problem. Where are remittances going? Where is money being sent out of the country? Where is money being sent within the country that doesn't make sense, not just from a pattern perspective but from a target perspective as well?

Sarabjeet Singh:

Absolutely, Greg. Risk will not sit in silos. It is the connections that reveal the real exposure for these institutions. What becomes critical is ensuring that the risk network is detected and brought to the surface.

That is where a unified data set across PEPs, sanctions and watchlists becomes relevant. You want to make fast, consistent decisions, while also addressing duplications, gaps and the number of false positives created by non-entity-resolved data. That is a significant issue for the industry.

As a result, workloads increase and budgets go haywire. Community banks and credit unions are already stretched in terms of the teams and budgets they have. If bad data is added to that workload, it can explode, leave risk unaddressed and expose banks to failure even more.

Greg Pinn:

I couldn't agree more. An integrated data approach is critical. Getting one part of an individual or entity's risk does not give you the full picture.

Understanding that someone is a PEP is useful, but understanding the context matters, too. Are they involved in high-risk activities? Have they been indicted or convicted? Are they both a PEP and sanctioned? That will open up your risk profile.

Some people are simply PEPs, and you need to understand that and handle it appropriately. Then there are nefarious PEPs, and there are sanctioned PEPs. Getting that full view is critical.

Financial institutions cannot do business with anyone on a sanctioned list. That does not just mean someone whose name appears on a sanctioned list; it means the people themselves. OFAC and other sanctions bodies provide some data, but integrating data across sources helps ensure that you are truly compliant.

This is what regulators increasingly expect. They want to know what steps a local financial institution is taking that regional, national and international banks have been taking for a decade.

The regulatory environment in the United States continues to be in flux. As we move toward a lighter regulatory framework, the pendulum can always swing the other way. As bankers and financial crime professionals, we need to be ready for what comes tomorrow and next year and plan ahead for how we do our screening.

Sarabjeet Singh:

Sorry, I was just going to add something specific to what we do with our PEP and watchlist data, and with sanctions data as well.

Typically, regulators publish a list for a country. They may provide a name and perhaps one identifier, which is very limited from an entity-resolution perspective. You are left trying to figure out whether one John Doe is the same John Doe who appears somewhere else.

We have gone to individual countries and identified additional data sets, all of them government-owned, to generate information such as an address or a tax ID. That information may also include the ID provided by the regulator. We use it to enrich the profile, which allows better unified profiles to be brought forward.

This is significant because the number of records we have deduplicated and entity-resolved has been substantial over time. That work is now reflected in the false-positive rates that have fallen for our clients. In tests clients conduct against other leading data sets, we are seeing a clear difference in false-positive reduction with these unified profiles.

So, Kate, to your question, this has a significant financial impact on budgets and teams, including the workload they handle. It is critical from that point of view as well.

Kate Randazzo:

Absolutely. And I'm glad you said that, Sarabjeet, because I think we're going to pivot to how we can use AI in this process. It is critical to have complete data sets and clear profiles, and to make fast, consistent decisions.

We know that one of the best ways to do those things today is to use AI, but banks have to be able to trust it. Sarabjeet, could you define the concept of a trust layer for AI? What does that look like in practice for financial institutions?

Sarabjeet Singh:

For us, when we talk about the trust layer, we define it as traceability, auditability, a clear audit trail and a validated data set. Those elements are critical for the data going into AI so that you can justify a decision.

The decision cannot be based only on predictions and statistical conclusions. It needs to be based on an actual justification using validated data. If the data set used by AI is one where the output links back to validated source data, human oversight is part of the process and audit trails are in place, all of those elements affect data quality and how much trust you can place in the data. That becomes critical in the AI world.

Greg Pinn:

One part of our approach to AI is asking how we trace the result back to its source, understand where the data came from, and model and evaluate the models we are using.

You cannot outsource your risk to another company, and you certainly cannot outsource your risk to AI. Saying, "The machine told me this was okay" or "The machine told me it wasn't" is not defensible.

That said, we've used technology in this space for decades. At its core, name matching is a computational decision about what will return a result and what will not. The alternative is printing out all the list data and reviewing it by hand. Nobody is going to do that, and it would be incredibly error-prone if you did.

This is the next tool and the next way to approach how you manage what humans are good at and what computers are good at. We sometimes think humans are better at these things than computers, but that is often not the case.

A great example is reviewing huge numbers of false positives. People are really bad at that. If you're hungry before lunch, upset with the person on the other side of the cubicle wall, or you've had a fight with your child that morning, you're going to miss hits. Algorithmic approaches are much more likely to raise the right items to your attention so that you do not miss them.

AI is another tool, another arrow in your quiver, that can help identify what your eyes need to focus on. It is not about removing human oversight. It is about determining where you want people with intuition, training, knowledge and experience to focus, and what information you want them to focus on instead of the tasks computers simply do better.

That is the balance we figured out with algorithmic approaches 10 or 20 years ago, and now we're figuring it out together as a community with AI.

Sarabjeet Singh:

To add to Greg's point, the PEP, sanctions and watchlist data we work with is public-source data. It is available publicly, and that is what is out there.

There is a thought process that asks: Can I search on my own? Can an agent do it for me? Can I write a prompt that is good enough for an agent to generate the right result?

We've seen enough situations where that approach does not return the right results. There are significant issues with entity resolution, access to the right and relevant data, and the searches that need to take place. The terminology used to write the prompt is another factor.

All of those components make it difficult for an agent to fetch information and return it with enough confidence for you to say, "I was compliant." Having a validated layer of data is critical in the AI world as well.

Kate Randazzo:

That makes sense. You mentioned access to data and the importance of writing the right prompt. What are some other common pitfalls organizations face when they are trying to operationalize AI while staying compliant?

Greg Pinn:

The biggest pitfall is using solutions that are not built for this purpose. General-purpose AI tools are built for general purposes. Even within those tools, some of us have experience with using Google Gemini when we want to make silly pictures, ChatGPT when we want to do research, and Claude when we want to write code. We become more comfortable with different tools being better at different things, even when they are all general-purpose tools.

In this space, the guardrails and prompt engineering built into custom tools are critical. You need someone who has experimented with and evaluated them.

When we talk about AI today, especially in mid-2026, we focus heavily on large language models such as Gemini, Claude, ChatGPT and Grok. We are not thinking as much about the historical AI and machine learning that has been built into many of these tools for decades.

That includes identifying patterns in name matching so you can better predict gender matches, predicting nicknames you have not seen before based on patterns in each language, and consolidating or resolving entities based on experience and training data within a data set.

A lot of machine learning has gone into the tools we use every day. Those systems have been tested over many years, so we've become comfortable with them. Now, to use a Spinal Tap reference, we've turned it up to 11, and we're all asking what AI can do for us and how we want to use different tools.

That is where you need to ask whether these tools have a trust layer embedded in them and whether they are built for purpose. We cannot rely on front-line AML analysts or even front-line AML managers to be prompt-engineering experts and make sure they are prompting appropriately.

Tools that synthesize and consolidate information for an analyst need to be built specifically with that understanding. They need guardrails to make sure they are presenting the right information, but they may also need certain guardrails turned off.

As AML experts and BSA officers, we spend much of our time researching human trafficking, sexual exploitation, terrorism, terrorist financing and organized crime. We are on the crime-fighting side, not the crime-committing side. Many general-purpose tools have guardrails that limit the ability to extract that kind of information.

If you try to find information on certain terrorist networks or patterns that help identify human trafficking, many general-purpose tools will stop you because they are designed to protect people. For a general-purpose tool, that makes perfect sense. In our world, we need tools with some of those guardrails turned off because we need to conduct that research, along with other guardrails that are enhanced.

It is always a balance between general-purpose tools and specific tools built for purpose.

Sarabjeet Singh:

Absolutely, Greg. I couldn't agree more about the risk of using general-purpose tools as a plug-and-play solution. That would definitely be a pitfall from an AI-adoption perspective.

In the diligence reports analysts complete, including KYC reports for client onboarding and EDD reports, the context layer is critical. As a data company, we understand the importance of a trusted data layer that gives these tools the right context and information, already cleaned to a certain level.

The tools should be looking at that information rather than everything in total. That is critical for adopting AI properly. Governance is also important: what information you use, who owns the data and how much you can use it.

There is a lot of unaffiliated data on forums that you may not want to use as a compliance or risk professional. Rules about what you use, who owns the data and how much you can use are all important.

The entire system is critical to the AI system you put in place, and human oversight remains critical. Yes, there is automation, but human oversight is critical in this adoption journey as well.

Greg Pinn:

You touched on something that is key here: data quality. That includes quality on the risk-data side, but also on the customer and transaction-data side.

AI can do a lot, but it cannot fix your source data. Making sure your house is in order makes a major difference in reducing workloads. You need a single source of truth within your CIS database and customer records. You also need to make sure you are funneling data appropriately and that your tools are set up correctly to move data from one system to another using the correct schema.

AI is great here, but let's make sure we handle the fundamentals first as well.

Kate Randazzo:

That's great advice. You touched on what banks should be doing to prepare their data infrastructure, but I have one last question for both of you. What is another major piece of advice you would give a bank or credit union that is considering AI for its compliance processes for the first time? What is the number one thing they need to know?

Greg Pinn:

The big question is: What problem are you trying to solve? Where are the slowdowns in your processes? Where are things getting in the way of identifying more risk or making your people more effective?

AI is a tool. It is a great and powerful tool, but you cannot wave an AI wand and make your problems go away. You need to do the upfront work of asking what you are trying to solve and which AI tools are built to help solve it.

There are tools that may solve problems in ways you never thought possible. That is what is amazing about the world we live in today. But you need to look for the risk you are trying to mitigate and the efficiencies you are trying to gain.

More specifically, look for the things taking up your people's time without helping reduce risk to your business, identify high-risk individuals and entities, identify high-risk transactions, or identify fraud.

It is about what problem you are trying to solve, not how you can bring more AI into your institution. AI is a tool that can solve those problems, but first identify the problem and then identify the best solution for your institution.

Sarabjeet Singh:

The other point I would add is ensuring that AI has the right data to work on. You need clean, normalized and linked data sets with the right data lineage, auditability and governance framework so that the outputs, decision-making and workflows are compliant, defensible and regulator-ready.

These elements become critical when building the infrastructure for AI. They allow the right network analysis, not just screening. The principle we've grown up with, GIGO, or "garbage in, garbage out," still applies in the AI world.

If the quality of the data is suspect, the outputs will not be optimized. For me, that is a critical component of any new AI infrastructure.

Kate Randazzo:

That makes sense. Thank you both so much for your time and expertise on this topic. Listeners, please check out the show notes for more resources from Resolute and Abrigo. We're glad to be working together to help compliance teams fight financial crime quickly and efficiently. Thank you so much.

Greg Pinn:

Thank you.

Kate Randazzo:

You can find this and future episodes of Ahead of the Curve on Abrigo.com or on your favorite podcast app or platform. Search for Ahead of the Curve, a Bankers Podcast, or search for Abrigo, and subscribe so you don't miss future episodes. Thank you so much for listening. We'll be back with our next episode as soon as we can.