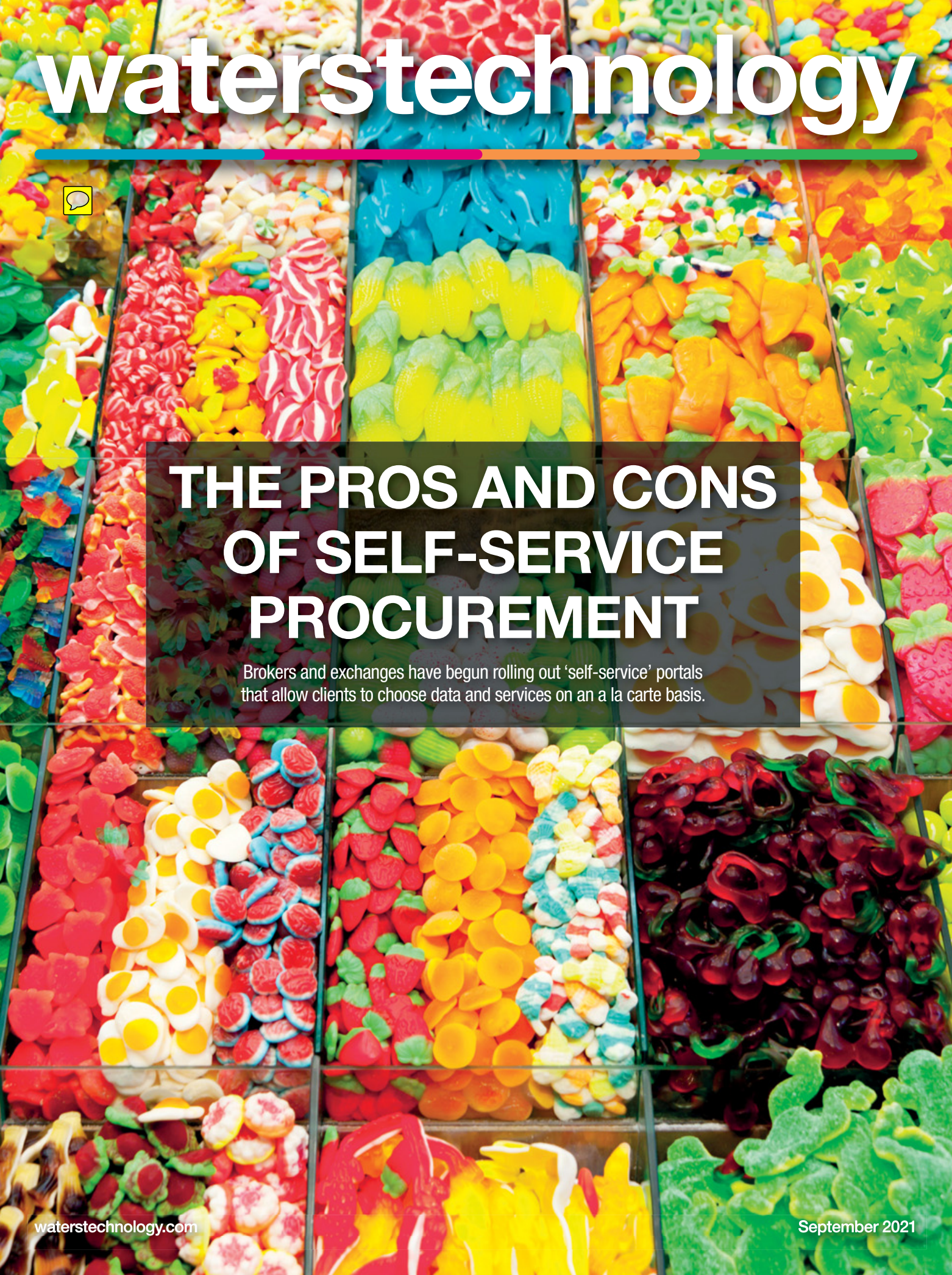
waterstechnology



THE PROS AND CONS OF SELF-SERVICE PROCUREMENT

Brokers and exchanges have begun rolling out 'self-service' portals that allow clients to choose data and services on an a la carte basis.

waterstechnology

Sign up to *WatersTechnology* newsletters

- Tailor your newsletter experience to suit your areas of interest
- Receive exclusive content handpicked by our editors
- Update your newsletter preferences at any time

Visit www.waterstechnology.com/newsletters to update your newsletter preferences today

Editor-in-Chief Anthony Malakian
anthony.malakian@infopro-digital.com
Editor at Large Max Bowie
max.bowie@infopro-digital.com
European Editor Jo Wright
joanna.wright@infopro-digital.com
Editor, Asia Wei-Shen Wong
wei-shen.wong@infopro-digital.com
Senior Reporter Josephine Gallagher
josephine.gallagher@infopro-digital.com
European Reporter Hamad Ali
hamad.ali@infopro-digital.com
US Staff Writer Rebecca Natale
rebecca.natale@infopro-digital.com
Reporter Nyela Graham
nyela.graham@infopro-digital.com
Head of Editorial Operations Elina Patler
elina.patler@infopro-digital.com

Global Content Director Victor Anderson
victor.anderson@infopro-digital.com
+44 (0) 207 316 9090
Commercial Director Ince Saleem
Tel: +44 (0) 20 7316 9258
ince.saleem@infopro-digital.com
Business Development Executive Sonja Patillo
Tel: +1 212 776 8083
sonja.patillo@infopro-digital.com
Account Manager Daniel De-Bruce
Tel: +44 (0) 20 7316 9126
daniel.debruce@infopro-digital.com

Marketing Manager Louise Sheppey
tel: +44 (0) 20 7316 9476
louise.sheppey@infopro-digital.com
Design Lisa Ling

Corporate and Single Subscriptions
US: Barbara Faiman tel +1 646 736 1852
info@waterstechnology.com

Global Brand Director Katie Palisoul
katie.palisoul@infopro-digital.com
Global Editorial Director Duncan Wood
duncan.wood@infopro-digital.com
Managing Director David Pagliaro
david.pagliaro@infopro-digital.com

Infopro Digital Head Office
Fifth Floor
133 Houndsditch
London EC3A 7BX
United Kingdom
tel: +44 (0) 20 7316 9000
fax: +44 (0) 20 7930 2238

Infopro Digital US
55 Broad Street, 22nd Floor
New York, NY 10004
tel: +1 646 736 1888

Infopro Digital Asia
Unit 1704-05
Berkshire House, Taikoo Place
25 Westlands Road
Quarry Bay
Hong Kong
tel: +852 3411 4888

Infopro Digital Customer Services
tel: +44 (0) 1858 438 800

WatersTechnology (ISSN 1068-5863) is published monthly (12 times a year) by Infopro Digital Risk Limited. Printed in the UK by Stephens & George Print Group, Dowlais, Merthyr Tydfil, Wales.

Reimagining data spend

I think this is a story that many of you will be familiar with: I hated cable TV and welcomed the advent of streaming services and the freedom they provided. I could finally pay for only the channels and services that I cared most about. Flash forward three years, and I have no clue how much I pay each month or year for the many streaming services I'm signed up for—and it could very well add up to more than my cable bill ever did.

I was reminded of this while reading Max's piece about the market data procurement process (*see page 22*). As he explains, several brokers (and even some exchanges) are developing so-called "self-service" data portals. As end-users increasingly want access to larger and more diverse datasets as they expand into new asset classes and geographies, they also want easier access to this necessary trade information. And these portals allow them to get the data they need, whenever they need it.

It's important to note that brokers are being pushed to evolve, though it won't happen overnight. "We are definitely seeing more of these data marketplaces, but I do feel we are a way off from seeing the true impact," Bernardo Santiago, founder of market data consultancy S4 Market Data, told Max.

A reason for this slow-burn shift is the lingering question of who gets access to these portals. Just as I lose track of how many streaming services I'm subscribed to and how many channels overlap, end-users face the same dilemma: There will need to be strict controls in place to prevent these portals from becoming a free-for-all with spiraling data costs and little way of tracking usage and spend.

But if nirvana is anyone being able to get the data they need, whenever they need it, do such controls undermine that ideal? If that's the case, there needs to be a wholesale re-think of how commercial terms and contracts are structured. For example, rather than using a traditional per-person agreement, terms would essentially shift to an all-you-can-eat model. Or, maybe it's a pay-as-you-go model, where firms assign each user their own budget, rather than centrally manage what content each user has access to. They can buy what they want within that limitation, but they have to manage their own budget. And once they hit the limit, there's no "data overdraft"—meaning they can't buy more packages until their budget is reviewed or refreshed.

These self-service portals could bring about a whole new way of looking at data spend. wt

Anthony Malakian
Editor-in-Chief

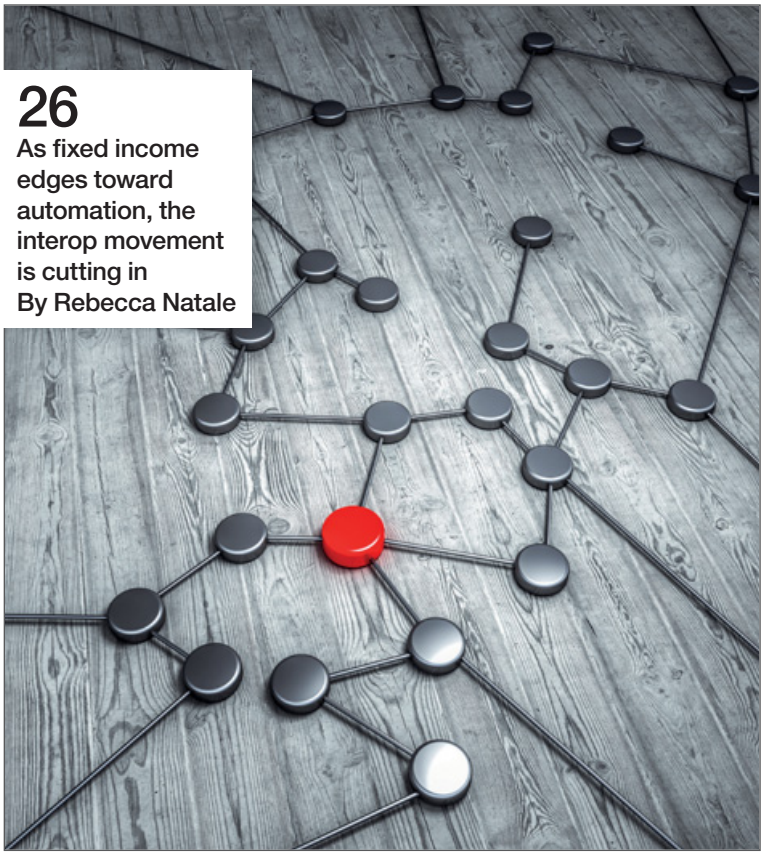
REMEMBERING OUR COLLEAGUES

It's hard to believe it's been 20 years since we lost 16 of our team members in the September 11, 2001, terrorist attacks. To honor their memories, *WatersTechnology* is speaking with family members, friends and colleagues and will publish a comprehensive tribute on our website, waterstechnology.com, in September. We will never forget them.

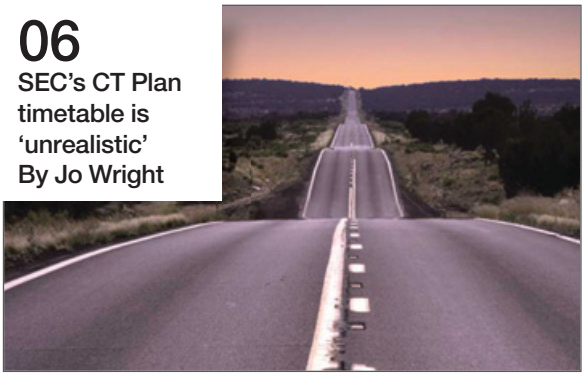
Contents

September 2021

GET MORE ONLINE
News. Webcasts. Video. Audio. Special Reports. Get it all at waterstechnology.com



26
As fixed income edges toward automation, the interop movement is cutting in
By Rebecca Natale



06
SEC's CT Plan timetable is 'unrealistic'
By Jo Wright

40
Jo Wright: SEC's market data plans sow confusion

30
Quants turn to machine learning to unlock private data
By Rob Mannix

UnaVista SFTR closure casts doubt on viability of reg reporting

The decision to shutter the service is another blow to the industry and the business case of reg reporting under SFTR. By Josephine Gallagher

In late July, UnaVista, the regulatory reporting arm of London Stock Exchange Group, announced plans to close its Securities Financing Transactions Regulation (SFTR) service on January 31, 2022, making it the industry's latest casualty in a battle with high running costs and wafer-thin margins.

The past two years have seen multiple providers in the regulatory reporting space roll back or offload their services—most notably Deutsche Börse's sale of its Regulatory Reporting Hub to MarketAxess, and CME's Nex Abide unwind—due to difficulties in remaining competitive and turning a profit from these entities.

The shuttering of Unavista's SFTR business has led some industry participants to question whether providing services around this regulation is economically sustainable.

Tom Wieczorek, head of global product management at UnaVista, tells *WatersTechnology* that following an annual review, LSEG decided to refocus efforts on its more successful regulatory reporting products, including those that cover the European Market Infrastructure Regulation (Emir) and the Markets in Financial Instruments Regulation (Mifir).

"Despite having quite a competitive offering within the space, we simply didn't get enough market share of what ended up being a smaller market than the industry expected to make a difference. So we decided to use those resources to make the offerings where we are really leading the market. It's just a reshuffling of our resources within the portfolio of products and services that we provide," Wieczorek says.

Sources at regulatory reporting vendors say that prior to SFTR taking effect in July 2020, volumes of reported transactions were much lower than expected. In the lead-up to the implementation, Wieczorek says, many vendors were confused about whether regulated entities would be expected to report intraday or end of day. Vendors calculated prospective revenues based on expected intraday reporting, but SFTR as implemented requires only end-of-day reporting.

John Kernan, chief executive of Regis-TR UK, says: "After several years of going to conferences where the International Securities Lending Association or the International Capital Markets Association spoke on SFTR many times, everybody was expecting many multiples of what we've actually seen when reporting went live."

As UnaVista's pricing model is calculated per message, Wieczorek says, lower volumes meant the company would have had to significantly raise its fees in order to continue running the service, a decision that would have rendered the product less competitive in a market that is already battling thin margins.

As *WatersTechnology* has documented over the past two years, a consensus among industry sources is that regulatory reporting is an increasingly difficult business in which to thrive, especially for regulated entities such as trade repositories (TRs), owing to their commercial limitations. Under Emir, all pricing for TR services must be made publicly available and approved by the European Securities and Markets Authority (Esma).

Kernan says TRs are not only hamstrung by the restrictions that come with being a licensed entity, they must also compete with unregulated intermediaries that have more commercial freedom. Unlike third-party vendors, for instance, TR's Emir requirements render them unable to develop new services with the data they collect or to cross-subsidize product lines.

The roadmap

Emir portability guidelines set out by Esma and adopted by the UK's Financial Conduct Authority (FCA) specify that TRs must provide the regulator and all TR participants with at least six months' notice of an intention to shut down the service. Wieczorek says LSEG had to alert the industry now rather than later to avoid clashing with other upcoming changes

to SFTR validation rules due to come into force after January 31, 2022.

"We have to do it now so that we have the six-month timeline that Esma mandates, and so that period finishes prior to the technical standard changes coming on February 1. This way, firms can move and work with their new trade repository on the new technical standards. We felt this would be the best way for the reporting firms to make the switch," he says.

LSEG first notified its clients of the closure plans in the last week of July. By that point, Wieczorek says it had already discussed the decision with both Esma and the FCA. However, many regulatory businesses spoken to for this story by *WatersTechnology* say they first learned of the decision to close the SFTR unit from media reports published on July 30.

The portability guidelines are intended to ensure that no reporting data is compromised or lost when firms switch from one TR to another. Wieczorek says LSEG learned a lot from last year's CME wind-down, and will follow Esma's standard process for migrating the data to a new provider while also providing clients with material to help them choose vendors and TRs. The information will include comparisons between UnaVista's services and those of other providers to help clients carry out due diligence when deciding where to go.

Once a client makes its decision, UnaVista, the new TR, the client, and any other third parties will need to work together to first transfer all open trades to the new TR, followed by historical trades.

Wieczorek expects UnaVista and its counterparties to be able to start the transfer process by mid-September.

The UnaVista SFTR wind-down will be on a much smaller scale than that of CME, which shut down multiple regulatory reporting businesses for Emir, including several TRs. SFTR reporting only began in mid-2020, while counterparties to over-the-counter derivatives trades have been reporting to Emir since 2012. There's simply been less time for a large volume of historical reports to accumulate.

> "Despite having quite a competitive offering within the space, we simply didn't get enough market share of what ended up being a smaller market than the industry expected."
> **Tom Wieczorek, UnaVista**

"It will be less disruptive than a CME switchover, as they were covering a few different regimes, such as Mifir, Emir, and others, so the scope was wider than here," says Ronen Kertis, head of global regulatory reporting at IHS Markit.

On the other hand, Kernan says, late movers could risk the same fate as those that took too long to move during the CME wind-down. An Esma report published in April 2021 says several counterparties were slow to select a new provider during the CME transfer window and regulators had to intervene to speed up the selection process.

The future state

UnaVista's departure from the SFTR business, plus the mounting cost pressures and lower-than-anticipated revenues, has caused some to question the viability of the SFTR regime itself. Several sources spoken to for this story say the regulation is economically unviable as it exists today, due to the high operating costs and low return on revenue. Some fear that other providers, particularly the remaining TRs registered for SFTR reporting in Europe—the DTCC, Regis-TR, and KDPW in Poland—could decide to follow suit and exit the market.

"As more and more vendors stop providing reporting services to financial firms, as we have seen with CME, Deutsche Börse, and most recently LSEG and their SFTR trade repository, choice across the market will diminish and new entrants are unlikely to emerge if there's a low commercial incentive to do so. This is certainly what we are seeing. As a result, it might become increasingly difficult for the regulator to enforce compliance," says Matt Smith, CEO of compliance technology provider SteelEye.

For several years, there has been debate over whether a single TR would be better suited for Emir and SFTR, rather than the current situation where several firms compete for business but operate under strict commercial conditions.

Kernan says competition is good for innovation, despite the limitations inherent in running a TR.

Kertis doesn't believe that a single TR will emerge, but admits he also failed to predict some of the changes the industry has witnessed over the past two years. "The existing players are robust and, as far as I know, it is not an insignificant business for them. So I think there will be stability. But, having said this, business is business, and sometimes decisions are made for reasons that are unrelated to that particular business," he adds.

SteelEye's Smith says that if a utility TR were to arise, it would have to be run by regulators. A potential concern is that a single utility operated by a business would have a monopoly on reporting fees and little incentive to improve the service.

Wieczorek says he sees TRs as an extension of the regulator, as in jurisdictions such as Switzerland the regulator nominates a provider—SIX in the case of the Swiss domestic market —to collect the data on its behalf. Whether this is a plausible option for the European or UK market is unknown. Some argue that competition is necessary, others say the regulation is unsustainable, but most agree that the commercial environment for regulatory reporting businesses is crying out for change.

What that change might look like, however, is unclear. wt

SEC's CT Plan timetable is 'unrealistic'

Implementing governance structure for new US public equities datafeeds within a year is highly unlikely, say industry observers. By Joanna Wright

The Securities and Exchange Commission has approved a new blueprint for the governance of the consolidated tapes produced for US equities, together with an aggressive timetable for its implementation. But industry observers say it is highly unlikely that full implementation of the plan is achievable within the regulator's 12-month deadline.

On August 6, the SEC quietly approved a modified version of what it has dubbed the CT Plan, which effectively consolidates the three plans that currently govern the two Securities Information Processors, the US public feeds for exchange-listed equities market data.

The SEC ordered the exchanges and the Financial Industry Regulatory Authority—collectively known as the self-regulatory organizations, or SROs—to come up with this plan, which forms part of its wider efforts to modernize the Sips. Perhaps the most important of the SEC's August 6 modifications was the inclusion of an implementation timetable for the plan, which was lacking in its draft form but demanded by market data consumers.

"So, you'll just have the one plan, but you still have three underlying subsystems that, presumably over time, become a unified system. That doesn't happen in 12 months—there's just no way on God's green earth," says Shane Swanson, senior analyst in market structure and technology at Greenwich Associates.

The SEC's timetable says the CT Plan must be up and running a year from its effective date, on August 6, 2022. In the approved plan, the regulator also identified actions that need to be completed in the intervening year and assigned deadlines to them, too.

> "That doesn't happen in 12 months—there's just no way on God's green earth."
> **Shane Swanson, Greenwich Associates**

"The most important thing that came out of this [modified plan] were the timelines. They're quite aggressive in that the effective date is the date of the commission's approval, which was August 6. So, the staged implementation starts from that date, and 12 months on from then, the LLC must be operative," Swanson says.

The CT Plan will take the form of a limited liability company (LLC) registered in Delaware. Within two months of the effective date, the LLC must form an operating committee to govern it, comprised of a balance of SRO and non-SRO voting representatives, including market participants.

"The commission went to great lengths to ensure that non-SRO voting members would have a meaningful role as full members of the operating committee. Several provisions of the plan that were formerly the decision of just the SROs are now the decision of the full operating committee," says Manisha Kimmel, chief policy officer at market data infrastructure provider MayStreet. Kimmel was formerly a senior policy advisor at the SEC.

Once the committee is formed, it also must adhere to a series of deadlines. For one, it must set the fees it will charge subscribers to Sip data. Businesses like MayStreet will be watching out for this fee schedule so that they can build their business cases and offer services as competing consolidators under the SEC's related infrastructure rule.

The operating committee must also enter into agreements with the processors operating under the existing plans—NYSE's Securities Industry Automation Corporation and Nasdaq—and contract a new independent administrator. It must also approve all necessary policies and

procedures to facilitate the implementation of the CT Plan.

These actions will all involve non-SRO voting representatives, Kimmel says.

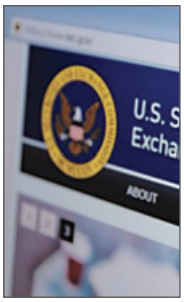
The SROs are required to provide regular updates on the implementation's progress. "The SROs also have to put out quarterly public reports on the plan, so everybody in the industry gets to see where we are as we go through the process. I do think that's healthy: sunshine is the best disinfectant. And if there are issues or concerns, they'll come up quickly," Swanson says.

Tall order

It's difficult to believe, however, that the SROs and, later, the operating committee, can accomplish all their mandated tasks in only a year, Swanson says. "The commission is saying this is going to go live within a year, and here are all the pieces that must converge for that to happen. I think the real question on everyone's mind is still, 'When does it actually occur?'"

He continues: "Even with the reasonable assumption that the SROs do everything they can to put this on the front burner, and put all their resources together: they get the operating committee selected; they get the fee schedule filed; the SEC then acts as quickly as it needs to with all the comment letters that will come in—and, trust me, there's going to be a boatload—the SROs select the vendor contract and the administrator; they have everything developed and tested—all the documentation, the billing, the audit function, and the policies and procedures are in place …. Even if all this happens smoothly, can that be ready to go 12 months from last Friday? And even if it can, is the industry ready for it?"

It's not clear how disruptive these changes will be to market data consumers, or whether the Sips' underlying technology will be different, Swanson adds. "Or is it just the billing practices that are going to change? As opposed to getting billed from the three separate entities that you have today, you'll get billed by one, but it'll maybe have three line items. What are the intricacies of that?"



The SEC says the plan is a crucial part of its Sip modernization intiative

Deadline tussle

During the period from October 2020 to August 6 this year, when the plan was subject to a notice-and-comment period, industry representatives said hard deadlines were necessary to ensure the CT Plan wasn't delayed indefinitely. Some commenters said that without the incentives of such deadlines, the SROs would have no reason to implement the plan at all. The larger exchanges view the SEC's modernization efforts as an existential threat to their businesses, and as unnecessary, since, they say, the current Sips perform extremely well.

In their comment letters, like the one NYSE submitted in November, the big exchanges said many of the plan's details remain too unclear for deadlines to be attached to them, particularly tight ones. There is still substantial confusion about how the CT Plan will interact with the infrastructure rule, they wrote. Nor is it clear how the proposed plan could include deadlines when so many unregulated entities must agree to provide services in order for it to become operational.

Comment letters said it would be impossible, for instance, to find and contract a plan administrator within the SEC's timetable. Some pointed to the difficulties in finding an administrator during the Consolidated Audit Trail project.

However, the SEC said that it "does not view the circumstances to be analogous. In the case of the Cat National Markets System Plan, the SROs were tasked with implementing the first-ever consolidated audit trail for equities trading, a complex NMS system without precedent. Here, by contrast, the operating committee will be conducting a request-for-proposal process to select an administrator to perform functions with which market participants, whether SROs or market data consumers, have extensive familiarity."

The commission insisted that, overall, its timetable doesn't underestimate the time needed to implement the CT Plan, adding that it wants the plan up and running as soon as possible since it is a critical piece of the Sip modernization initiative.

But the SEC doesn't spell out that it will refuse to grant extensions to the CT Plan's implementation deadline. In analogous regulatory actions, the regulator has explicitly stated that such extensions would not be granted.

Legal threat

The exchanges have already sued the SEC over the plan and the infrastructure proposal. In oral arguments in May, NYSE and Nasdaq argued that the regulator was overstepping its authority with the order that led to the creation of the CT Plan. The court ruled in June that it was too early for the exchanges to challenge the proposal, as it was not yet a final action that it could review.

Industry sources would not be drawn on whether more legal action will disrupt the SEC's timeline for implementing the CT Plan, but they believe such legal action is likely. In a blog post, MayStreet's Kimmel noted that the exchanges re-filed their lawsuit in April, anticipating the dismissal.

Swanson says: "I would assume the SROs will ask for some sort of injunctive relief. Whether or not they get that is another matter. They must show the likelihood of success on the merits [of their case], and that is a pretty high burden of proof. They haven't been as successful in this round of things as they were previously."

In 2020, the large exchanges won a legal battle with the SEC over market data fees.

"It's a tricky question, because the SEC has proposed all of this as modifications to the CT Plan, and the SROs are arguing that some of these changes are outside the SEC's authority," Swanson says. wt

Trading Technologies adjusts development strategy as it awaits new owner

The futures trading platform recently rolled out a new OMS offering, while other projects, like its Echo Chamber market data platform, have been put on pause until a sale goes through. By Anthony Malakian

The rumors that Trading Technologies (TT) was set to be acquired began last year, and the reports pegged Goldman Sachs as the leading candidate to buy the futures trading platform provider—and then those talks fell apart.

Even with that setback, the vendor, which is partly owned by Chicago-based investment firm The FSB Companies, is still very much on the block to be sold—and sources have indicated that a sale will likely happen soon, though the buyer's identity so far remains a mystery. This puts TT in a slightly awkward position; the company can't make any big, flashy investments, and some of its more ambitious projects have been put on hold to ensure that these investments align with the new owner's roadmap.

Take, for example, the rollout of its new order management system (OMS), which went live last fall. Farley Owens, TT's recently appointed president, views this offering as being "probably our biggest area of growth potential," but some clients are on the sidelines, waiting to see who the eventual owner will be.

> "We've actually signed three deals that have yet to be announced across three different types of customer. So that's three different types of firm utilizing our OMS in different ways."
> **Farley Owens, Trading Technologies**

"Some of our customers are waiting to see what happens with the sale of TT," Owens says. "They want to know who's going to own TT before they make a commitment to spend more money and do more integration—if it's somebody they don't like, they don't want to do that."

But even as some existing customers wait, TT has won over a handful of converts. "We've actually signed three deals that have yet to be announced across three different types of customer," he says. "One of them is a global brokerage firm that does execution for buy-side firms; we signed a deal with a pretty sizable regional bank; and then we also signed a deal with a commercial energy firm. So that's three different types of firm utilizing our OMS in different ways."

The OMS came about thanks to a decision made several years ago to sunset the vendor's flagship X_Trader execution management system (EMS)—a monolithic, mostly closed-off system that was designed in the mid-1990s using a client-server architecture that was initially made available only for on-premise deployment.

Today, Trading Technologies has migrated almost all of its users from X_Trader to the new TT trading platform. "We literally have one customer with one desk accessing one market," says Owens, who has been with the company for almost 20 years and took over as president in February. "It's one of our largest bank customers that still has access to Euronext while they work out their back-end integration for wholesale trades. So it's one or two users that will still be on the platform for another couple of weeks. Everybody else is on the TT platform."

Echoes of the past

The futures trading platform space is a competitive one that has seen a lot of consolidation. Ion Group has become a powerhouse there, like it also has

in fixed income—especially after it acquired Fidessa, though that integration hasn't been without its issues. Broadridge has also grown its futures trading presence with the addition of Itiviti, and well before that deal there was FIS–SunGard. Mainly in Asia, CQG competes with TT, and for options, there's Vela Trading Systems, which recently merged with Exegy.

In this competitive environment that features some larger contenders, under the leadership of Rick Lane—the former CEO of Trading Technologies who left earlier this year to join Citadel as the CTO of core engineering—the company burnished a reputation as being at the leading edge of innovation in futures.

"They have always been innovative and when you look at the broader space, a lot of their competitors have been acquired," says Brad Bailey, research director with consultancy Celent's capital markets division. "The futures industry is an important industry and the products in the space are key for so many different people, and not just financial firms—it's the hedgers and the farmers, too."

While almost every vendor is now ditching legacy platforms in favor of cloud-based offerings, TT made that decision earlier than others. In addition to the development of an OMS, the company was also early in developing institutional-grade tools for crypto trading; it launched an infrastructure-as-a-service offering; it started plotting a move away from screens in 2018; and the year before that, it made a major investment in Neurensic to improve its machine-learning capabilities.

Perhaps most ambitious, though, was a project to help firms reduce their market data costs in futures trading. The platform, dubbed Echo Chamber, was designed to combine futures contracts traded on different exchanges to provide a single view of each, no matter where the contract is traded. (*See Echo Chamber's UI, right.*)

As is true of any inspired idea, the project needed significant buy-in from large trading participants in order for it to achieve its full potential as something of an industry utility that helps firms to cut their market data costs, rather than simply being an internal risk application. But because of the uncertainty created by the coming sale of TT, Echo Chamber has been put on hold, Owens says.

"As we were going through the sale process last year, the leading contender to buy us actually told us that they'd prefer if we did not work on that, and so we kind of put it on the shelf. It's there, and we still have some customers expressing interest to use that internally for their own trading within their own firm—like a large prop shop. And then there's the broader version of that, which is a more global feed that we would publish out," Owens says.

Owens says, however, that internally TT still believes in the platform, that "it's a viable product that people would be willing to pay for," and that it's something the vendor will "probably revisit as soon as this sales process is over."

Waiting…

Celent's Bailey says that it makes sense that there's interest out there in acquiring Trading Technologies, and that interest has to do with both historical trends and where there are still needs in the market.

"Futures trading has been electronic for quite a long time—and even options on futures are getting more electronic—so it's a competitive space, which makes TT interesting," he says. "Additionally, equity algos have grown much faster there than in futures. So even though you have systematic and automated futures trading, CTAs and CPOs and all those traders still might not be getting all they can from execution algos in futures."

Owens notes that while TT has had to ease off the accelerator on some projects, it's not as if its developers and engineers are standing around. In fact, he says that now that the migration from X_Trader to the TT platform is essentially complete, there's a long list of projects that need to be addressed.

"At this point, having just completed the migration, we have a long pipeline of things to do, whether that's outstanding customer requests, or new markets, or some of the OMS functionality that we have yet to deliver—we really have more work than resources at the moment," he says. "

And the instant there is clarity around the future of the company—because TT has invested in cloud and has opened the platform up to include support for an array of other markets and asset classes beyond futures and options—Owens says they will be able to adjust quickly.

"We think the new platform is flexible enough with the newer technology that we can shift gears and move quickly whether to support a new market, or integrate with other vendors in ways that provide unique solutions whether in data analytics or other things," he says. "So it will be a little bit of a wait-and-see, because depending on who does buy us, they will have their own things that they may want to integrate with us to provide solutions to their customers." wt

Echo Chamber's user interface

Charles River, Wave Labs team up for enhanced OEMS

The strategic partnership will involve a three-part integration including system connectivity, combined visualization and the creation of client feedback loops. By Josephine Gallagher

State Street's Charles River Development (CRD) is integrating its fixed-income order execution management system (OEMS) with start-up Wave Labs' eLisa credit execution platform. The move is in line with a broader shift among buy-side technology providers such as heavyweights like BlackRock Aladdin and SimCorp, which are opening their technology frameworks and interoperating with a variety of third parties (in some cases, rivals) to deliver vendor optionality to the asset management community.

Over the last few years, choice has become one of the biggest unique selling propositions for buy-side technology vendors marketing to the buy side. Michael Beattie, head of the product strategy at CRD, says that although partnerships were not part of the company's founding DNA, the vendor has changed that approach and made it central to its strategy since State Street's acquisition of the middle office business in July 2018.

"We want to be on the side of choice here, and have tried to embrace partnerships over the past two years. We want our clients to feel like they have a choice in the vendors that they interact with," he says.

Beattie recalls that, during his first interaction with Miles Kumaresan, founder and CEO of Wave Labs, it became clear they had a similar perspective on building fixed-income solutions. Rather than trying to retrofit an equity system to work for corporate bonds, they both thought that credit platforms should be specifically designed to meet the needs of the credit markets. The corporate bond market differs from the stock market because of the variety of bond types, which are further differentiated by characteristics like issuance date, maturity, and size.

> "We're trying to pick partners that have the biggest bang for their buck, where clients trust them as a provider, where clients have a relationship with them in some way and there's something unique about what they do that's helping to drive the process for us."
>
> **Michael Beattie, Charles River Development**

While CRD is now more open to third-party integrations, and there are many vendors the firm could work with, there are limits to the strategic push, Beattie says. It's not feasible to work with everyone because of timing, the workload involved, and because there can be significant competitive overlap. A big factor in choosing a partner, and what it boils down to, he says, is client demand and the value the integrated systems can offer the mutual client base.

"We're trying to pick partners that have the biggest bang for their buck, where clients trust them as a provider, where clients have a relationship with them in some way and there's something unique about what they do that's helping to drive the process for us. And frankly, it's got to be something that we aren't planning to build or that we think we just don't have the time to build. That's where this fits into the picture," Beattie says.

Beattie says that to justify working with a partner, there must be significant demand for the integration to occur. Traditionally, for CRD to move forward with a partnership, at least 10 clients of the vendor's 300-strong roster should ask for the pairing. In this case, Beattie says there were three major CRD clients asking for connectivity to eLisa, which participated in tri-party discussions to use the platform via CRD. In addition to those three, "several others" have expressed an interest in the pairing. Neither CRD nor Wave Labs could disclose the names of the three buy-side firms due to vendor-client confidentiality, but all were managing assets north of $350 billion.

"We [been having] tri-party conversations between me, Kumaresan, and the clients. That's been eye-

opening for us, because we understand the technology and what we're hoping this will be. When you bring a client into the mix, it does a few things: it validates our assumptions in the workflows and it ensures that there's something real there, that someone's going to use that system," Beattie says.

Carving out a niche

The eLisa credit trading system is designed to help traders identify liquidity in a fragmented market, and better visualize and execute trades. Wave Labs currently connects with multiple fixed-income trading venues—including UBS BondPort, MTS, and the Ice Bonds ecosystem—to aggregate liquidity via APIs onto one single user interface. In most cases today, investment firms manage multiple venue user interfaces or call records when trading corporate bonds—typically using an OMS or via requests for quote (RFQs).

WaveLabs is also in the process of connecting eLisa to Tradeweb and other venues, as well as sell-side firms, for streaming prices and sending RFQs. According to Wave Labs, the system is designed to receive 80,000 quotes per second and process 25,000 quotes in real time.

On the trading side, the credit execution system will be plugged into CRD's OEMS to facilitate bi-directional workflows between the portfolio managers and traders. In May, the start-up also announced its connection to IHS Markit's thinkFolio investment management platform, and is also in talks to connect with SimCorp. The eLisa system provides four levels of pre-trade market analytics that contextualize where a bond should be trading, where it has been trading, where it is currently trading, and predictions about where it should be trading. Kumaresan says the last set of analytics, the predictive pricing data, is almost nonexistent in the market today.

"We calculate the fair value based on the conditions in the market, how other bonds are trading, and then predict where that price would be at a certain future time," he says. "So, we are giving a trader a guide to discover where they can and should be getting the bond they are trading."

In its post-trade analysis analytics, the credit trading platform provides execution analysis, automated best execution documentation, and data on trading performance.

The system uses a proprietary machine learning algorithm designed to study the behavior of the trader and provide recommended trading strategies, price settings, and smart choice trading protocols. In other sets of analytics, the credit execution system also provides broker scoring used to evaluate reliable sources of pricing information and mitigate against information leakage. The scorecard would be used to safeguard against instances where brokers spoof prices and fish for information in the market.

"We provide pre-trade information of market color and analytics that remove the noise by calculating broker scores so that the trader only sees the brokers that are reliable," Kumaresan says.

The integration roadmap

Over the next year, the CRD and Wave Labs integration will be broken down into three parts: the initial testing and certification of the systems, ensuring they successfully operate together; the embedding of eLisa execution visualizations into CRD's user interface; and the creation of a systematic feedback loop to improve its user experience across the two platforms.

In the last two weeks, the two vendors completed the first phase of the integration, clearing technical hurdles that include sending FIX messages to and from CRD's OEMS and eLisa. Beattie says this is to ensure that price discovery and execution information can be carried out on eLisa and that the information is then accurately communicated back to the CRD's OEMS.

Over the next year, as part of the second phase of the implementation, CRD and Wave Labs will work to embed aspects of eLisa's visuals into CRD, so the two can operate as a single platform with instant updates between the CRD blotter and the credit execution system.

"What is most interesting is [eLisa's] visualization of how the orders are executing, alert information, and so on. A lot of that framework today does exist in Charles River, but what we envision is taking pieces of that, so that you don't have to jump between the different systems, and so that you've got one core platform, and you're looking at a combination of Charles River's order blotter with pieces of eLisa embedded into that," Beattie says.

In the last phase of the partnership, the two vendors aim to create a feedback loop that will help them better understand how their clients are using the platforms and how to improve their user experience. The feedback loop will be informed by eLisa's ML algos that are built to learn how traders interact with the platforms and help them to enhance their performance and automate trading functions. Beattie says the integrated systems could pull execution information from eLisa and other third parties to help inform the creation of orders in portfolio management and trading activity such as routing and knowing where to access reliable liquidity.

"We're headed toward providing clients with the choice that they want, understanding how they use our system—with them being part of the integration—and talking with clients," Beattie says. "[Using this approach] we create what we call 'client personas' of a trader and a portfolio manager, and understanding what they're doing with these systems is really important," he says.

The eLisa platform and the first phase of the integration will officially go live in October. wt

Not a drill: Alt data providers push oil and gas for investment

As the oil and gas industries say they're "greening" operations, vendors are offering ESG data and metrics to help investment professionals understand future risks and financial returns. By Hamad Ali

Some industries—such as concrete, and extractive industries like oil and gas drilling—are major polluters and contributors to climate change, but they're integral to human civilization. Investment professionals may opt to leave these companies out of portfolios that screen for or are tilted according to environmental, social, and governance (ESG) criteria. However, alternative data providers are offering data to help portfolio managers recognize the greening efforts of corporates in extractive sectors.

It's important to include these industries in some portfolios, says Blake Scott, president at alternative data provider Waste Analytics. While demand for oil is decreasing, the world is still heavily reliant on fossil fuels. Active investor engagement will encourage companies in these sectors to make genuine moves toward greening their technologies—particularly when it comes to how they deal with the waste they produce in their extractive practices.

"If you don't invest in them, you are going to create a situation that at some point, when we get to the place where oil and gas are not the main drivers of energy, these liabilities will be left to the general public to clean up, and from a personal perspective, I think that is unacceptable," Scott says.

Prior to his current role, Scott ran a company that recycled oil drilling waste. In 2018, he founded Waste Analytics, which focuses on data on the waste that oil and gas companies produce as they drill. The vendor is currently trialing its data product with asset managers.

Active investor engagement in fossil fuel firms could help their efforts to go green

Drilling for oil produces both fluid and solid waste, including salts, heavy metals, and wastewater, which can be disposed of in different ways—some more sustainable than others. Waste Analytics provides waste datasets gleaned from corporate reports and regulatory filings. Each company has its own level of disclosure, and Waste Analytics investigates whether the company has a sustainability report, and within that, if it collects relevant information on waste.

The oil and gas industries are considered controversial by some investors, due to decades-long histories of environmental disasters, unsafe extractive practices, drilling on indigenous land, and funding dictatorships. These sectors have made gestures towards greening their operations—switching to renewable energy, for example. Many of the world's biggest energy firms, including BP, TotalEnergies, and Shell, have pledged to achieve net-zero greenhouse gas emissions by 2050, the European Union deadline for carbon neutrality. The EU also recently proposed a ban on the sale of new petrol and diesel cars by 2035. All of this could have a cumulative effect on how investors approach the oil and gas industry.

To understand how these companies stack up against one another, and to be able to track their sustainability impact, analysts need data. Alternative data is one way to get a more in-depth look at the ESG performance of an industry that is under a lot of pressure to adopt more sustainable methods. For investors, the performance of the ESG data around these companies can also have an important impact on their financial performance.

Another alternative data provider, Sentifi, looks at Twitter, news, and blogs to find data about companies, commodities, and events, such as oil spills, that could impact corporates' valuations.

"We would capture the company that was responsible for the oil spill. We would look at the sentiment and attention shifts around that oil spill," says Sentifi CEO Marina Goche. "How is that oil spill being perceived for the company? How is that oil spill being perceived relative to peers who may have had oil spills? How has this particular oil spill been perceived relative to previous oil spills by this particular company? Is it likely to impact the asset price? We look at the sectors and industries that are impacted by this oil spill," Goche says.

The vendor has developed a machine learning model that analyzes the credibility of a source by looking at factors such as the source's identity, whether the content is unique or merely a retweet of another source, if it's a bot, and what types of followers the source has on Twitter.

In most cases, inaccurate or insufficient corporate reporting is about a lack of capacity or knowledge within an organization. In some cases, though, it's an outright massaging of the truth. Scott says that one US-based oil company, for example, states in its sustainability report that its disposal of drilling residue is tightly regulated. However, Scott says, its method is simply to bury the waste, which is not an environmentally friendly approach.

"I don't consider that to be highly regulated. Investors need to be able to understand this to see what's going on," he says. wt

‘Connect’ schemes will force Chinese buy-side IT overhaul

Initiatives like Stock Connect and Bond Connect, which provide greater access to international markets, will drive change at Chinese asset managers struggling with legacy trading tech. By Wei-Shen Wong

Today, onshore Chinese asset management firms maintain a plethora of systems, often cobbled together to bridge the shortcomings of other providers. But looming initiatives intended to drive greater participation by Chinese investors in international markets could stretch these setups to breaking point, software vendors warn.

The key problem is that there’s no true all-in-one system in China—nor is there a very complete concept of a portfolio management system, says Iris Wang, global head of strategy and head of China at Enfusion Systems. In July, the Chicago-based cloud investment management software and service provider obtained a wholly foreign-owned enterprise (WFOE) license in China, which allows it to start operations in Mainland China and hire locally.

“They have various legacy systems from different providers. The last time I counted, as an example, one of the bigger players has roughly 12 different systems. Some in-house, some third-party, all cobbled together,” Wang says.

Even if these disparate systems can connect to each other, they generally don’t communicate with each other well. “It’s not real-time; it’s certainly not efficient in terms of data flow and reporting out. And it’s definitely not accurate, because you have to do multiple reconciliations between different systems,” Wang says.

This legacy setup—while convoluted—may work just fine today for onshore managers’ domestic business, but increased access to international markets will put greater pressure on these systems and drive firms to seek more efficient replacements.

China currently lacks an all-in-one portfolio management system

Impetus for change

Chief among these is the introduction of “southbound” trading on the Bond Connect initiative—a cross-border trading program established in 2017 to open up China to more international investment via Hong Kong. The first phase of Bond Connect was “northbound” trading, which allowed foreign investors to access mainland China’s domestic bond markets. That went live in 2017.

Southbound trading is expected to go live later this year, and will allow mainland Chinese asset managers to participate in international bond markets. This is expected to result in far greater volumes of order flow and market data, which onshore asset managers’ systems may not be able to handle.

Northbound and southbound trading are two very different things, says Bing Li, head of Asia Pacific at Bloomberg in Hong Kong. “The Southbound Bond Connect is completely different in that China is now a large developing market coming into the international space, and trying to take in data and understand the technology, so the approach has to be different,” Li says.

Specifically, the China market that is investing offshore is limited to institutions like the large commercial banks, sovereign wealth funds, and those participating in qualified foreign institutional investor programs.

This will require different approaches to technology and connectivity. While the international markets are certainly no strangers to electronic trading, the onshore community, at this point, is still foreign to the concept of electronic trading and workflows that use request for quote, for example. “A lot is still done over voice,” Li says.

In contrast, from a data perspective, the introduction of Southbound Bond Connect will see international bond market data being consumed at scale by onshore Chinese buy-side firms for the first time. This—and anticipated increases in trading volumes—will create more demand for automation.

“The only thing that drives innovation is demand. If I’m sitting there as a market maker or a bond trader and I get five orders a day, I don’t need automation. More activity and volume will demand a reaction that demands more automation,” he adds.

While global providers like Bloomberg can serve clients with its various data products, it’s still early days for the bond trading infrastructure in China, and will require participation from local and international participants to develop the market.

For example, this may include international data providers adding more context to their data, to provide clarity for Chinese investors who are unfamiliar with overseas listings.

“If Home Depot is coming to market for new issuances, Chinese investors would ask, ‘Who is Home Depot?’. You normally wouldn’t need to provide that much information about Home Depot, as it is a recognizable brand name and its issuing and rating history is disclosed in a format that is familiar with most international investors,” he says. “But with this new group of participants in the international space, we need to be much more organized with how we present that information.” wt

Burden of Proof: The IEX breakaways looking to shake up broker algos

Founded with the principles of "transparency and academic rigor," some say Proof's model and technical approach is a test case for a new generation of cloud-native broker startups. By Max Bowie

Proof Trading, a startup agency broker-dealer founded by former IEX executives Daniel Aisen and Allison Bishop, is about to launch its second trading algorithm in a bid to deliver more efficient and lower-impact execution.

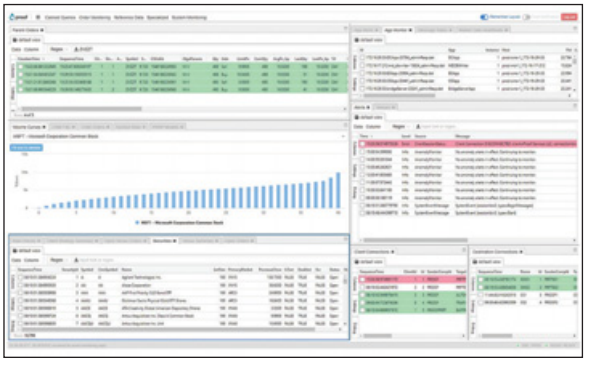
The firm's first algorithm, which went live in March, is a VWap (volume-weighted average price) execution algorithm. The second algorithm, expected to go live at the start of August, serves as both a liquidity-seeker and impact minimizer, which first seeks out large blocks of liquidity to execute large chunks of an order against. Once it has picked off large blocks—or if it can't find any—the algorithm steadily executes trades under the radar to avoid any large impact on price.

"We have historical models about how traditional styles of trading impact price movements, which gives us an idea of the differing levels of risk between trading styles and liquidity," says Bishop, who is president of Proof. "The result is that the algorithm is constantly reevaluating what is a good way to trade, and the schedule to trade on."

Reducing execution costs is a key element of Proof's proposition, yet that's not an easy thing to quantify. "It would be irresponsible to say, 'We're going to save $X in impact costs over other brokers,' and anyone who does say that is probably lying," Bishop says.

And that's not evading the question or taking a swipe at the competition: execution cost is a convoluted issue.

"The Street thinks of agency execution as a commodity—everything is low-touch, and all the brokers have their own, very similar, offerings," Aisen says. "Execution costs are already a big cost for firms, but they are very hard to measure because this area is a complete 'black box' and is very opaque. Once you have a better way to measure these costs, you can start to compare them."



An example of the Proof Trading interface

This "better way" draws on Aisen's and Bishop's backgrounds in quantitative research. Traditional transaction cost analysis is a relatively straightforward calculation based on measuring outcomes. The problem, Bishop says, is that traditional TCA lacks context, whereas execution performance can vary wildly based on external factors.

Any firm trying to measure and minimize the impact of its own trades must take into account myriad other factors, such as the movement of other stocks in the same sector, and the activity of other market participants.

"So how do you measure the amount of noise and give context for interpretation, and reduce that noise to improve execution? We do robustness checks and take out some of the data to see if the results are being influenced by a small number of outliers," she says. "By opening the black box and doing surgery on the prices, you can get higher-quality information compared to just looking at the outcomes."

Open and transparent

Aisen says that what sets Proof apart is its "extreme transparency and scientific rigor." Indeed, the two go hand-in-hand: Bishop notes that—unlike in finance, where every new development represents a competitive advantage and is kept a closely guarded secret—in science and academia, everything is published and open to peer review, a philosophy that Proof has tried to adopt.

"In the cryptography space [Bishop's chosen field], you expose encryption protocols to the scientific community over decades," she says. "So why not have the process of designing and measuring algorithms open and transparent, so you can crowd-source inputs and the buy side can directly assess and contribute to it?"

Brad Bailey, research director for capital markets at Celent, says Proof is "extremely transparent" about its data and the methodology of its algorithms and arrival prices, and it provides transparency into the data, logic and methods through which orders are handled on behalf of their clients.

"Are there firms that want this type of data? Yes! 'Show me the data'-type discussions are becoming ever-more prevalent," he says. "There's a growing need for people to have more detailed transparency. As things like plumbing and black boxes become more sophisticated, the need for greater transparency and understanding becomes more important. wt

Research management systems vie to double as data, analytics providers

RMS provider Sentieo feels the pressure to become quasi-data and analytics providers in their quest to cover the gamut of the buy-side research analyst workflow. By Rebecca Natale

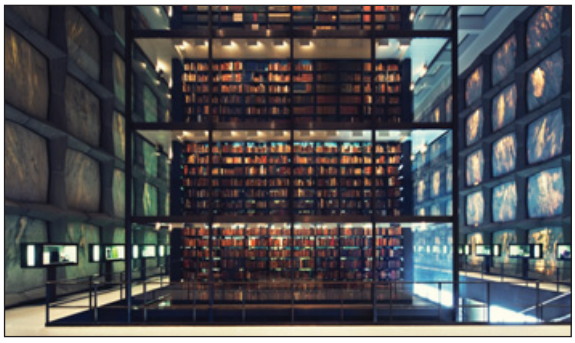
Research is a slog. Teams of analysts sit at their desks, fingers typing furiously, windows and tabs opening and closing. At least, in the pre-pandemic days, they had each other.

Though it's still unclear how many will work remotely indefinitely by choice, firms are grappling with the possibility that they could be forced out of their offices again—perhaps by another pandemic wave, or through displacement due to climate change-related events, which are ramping up in frequency and intensity around the world. In addition to hardships these scenarios cause on a humanitarian level, they also throw a wrench into everyday processes like collaboration, data-sharing, and workflow efficiency.

Hodges Capital, an investment advisory firm based in Texas with more than $500 million in assets under management, enlisted the help of Sentieo, which provides a cloud-based research management system (RMS) for investment analysts and researchers. Eric Marshall, director of research and co-chief investment officer at Hodges, says overhauling its proprietary research department with Sentieo has proved "transformational" for the team of 11 he oversees.

Having had a proprietary research process in place prior to the pandemic, Hodges' analysts published all of their research notes on an internal drive, which featured a "clumsy" interface to access, Marshall says. If a portfolio manager needed information quickly, it was often easier for them to find an analyst directly and ask their questions, rather than sift through the database.

And it didn't offer the research team a way to aggregate sell-side research from other institutions, meaning they had to seek out individual reports online or query their email inboxes. When the firm went remote due to Covid-19—though staff have since returned to its Dallas office—those issues were compounded by reduced bandwidth and screen real estate, as the number of necessary daily calls rose.



Sentieo wants to provide a place for analysts to perform their entire research workflow

Sentieo offered Hodges a way to pull data from industry journals, sell-side reports, internal research, and regulatory filings into a central hub that could be accessed and analyzed by multiple users at once, in the style of Google Drive.

More than an RMS

Though Sentieo is one of a handful of RMS providers—others include MackeyRMS, which recently merged with InsiderScore, as well as Nasdaq's eVestment, FactSet RMS and SS&C's Tamale RMS, to name a few—it views itself, to a certain extent, as a quasi-competitor of big data providers such as Bloomberg, FactSet, and S&P Global Capital IQ, says CEO David Lichtblau.

While institutional data providers offer critical functions like sell-side and exchange connectivity, news, and fundamentals, he says they don't provide a place for analysts to perform their entire research workflows—which includes pulling numbers together, getting them into Excel, modeling, comparing similar companies, writing notes (usually in Microsoft Word), and finally emailing them to a portfolio manager and storing them in Dropbox or SharePoint—in one single application.

This is the crux of Sentieo's RMS, but the company doesn't directly collect fundamentals from companies. To fill the data gap, it began licensing market data from S&P Global, Refinitiv, and FactSet, and alternative data from Earnest Research.

"We are aspects of the [Bloomberg] Terminal. We're not the next Bloomberg, but we are a platform to do the work. We've got market data, we have a built-in RMS, and then part of our secret sauce, if you will, is all the work we do around natural language processing," Lichtblau says.

A function called Search has collected and structured 50 million text documents—including filings, annual reports, financial news releases, and all broker research from Goldman Sachs and its peers—all of which Sentieo has indexed and applied financial linguistics to, in order to understand where, in long pieces of text, a speaker is giving guidance or identifying risk factors, without directly saying so.

"We still have Bloomberg—I wouldn't say Sentieo is a replacement for it—but there are functionalities that are just easier to do at Sentieo that we no longer use Bloomberg for," says Hodges' Marshall. wt

Dutch asset manager turns to decision trees for currency predictions

APG has improved prediction accuracy for G10 currency movements after adopting decision tree-based machine learning. By Hamad Ali

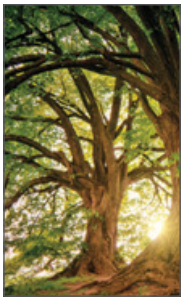
Decision trees are among the more popular applications of machine learning in the capital markets. Uses include finding patterns in request-for-quote (RFQ) datasets and predicting stock prices. Decision trees employ a tree-like model to map out decisions, or "nodes," and the probabilities of consequences branching out from these nodes, to work through a database and calculate outputs.

Amsterdam-based pension fund APG Asset Management, with $690 billion in assets under management, is using decision tree-based machine learning to predict the movement of G10 currencies.

"We are making predictions on currencies using market data, economic data, and fundamental data. For example, we are trying to predict if in a week or a month, how different currencies are likely to go up and down," says Luiz Felix, expert portfolio manager at APG.

Besides economic and market data, APG also uses predictions arrived at using nowcasting—predicting the present or the very near-term future—among the inputs into its machine learning model for forecasting future currency movements. He says his team uses principal component analysis (PCA) to streamline lots of daily economic data to use in nowcasting for different regions.

"The decision trees are what are called supervised learning methodologies. In the case of nowcasting, the PCA is an unsupervised learning methodology, which tries to find commonalities from lots of data. So, the PCA is quite a different animal from the tree models," Felix says.



APG is using decision tree-based machine learning to predict G10 currency movements

Nowcasting can help a researcher understand the dynamics of local economies, he says. "Nowcasting utilizes the evolving set of macro data available through the month and on a real-time basis, adjusts our understanding of the economic cycle in different countries."

It does this by using two sets of data. One is macroeconomic releases, such as the Purchasing Managers' Index, which is used to indicate economic trends in manufacturing and service industries; and inflation and employment figures. The second dataset is the expectations of economists about those figures.

"Economists are also changing their forecasts about those economic releases daily. And by tracking both the realized data and the forecasts, we produce our nowcasting," Felix says.

Analysts have to manipulate the data series that goes into the machine learning methodology. "You want to do data preparation to embed expert knowledge. You don't want to simply get raw data, throw it at the machine learning methodology, and hope that the algorithm will actually pick up a signal. You want to input data series that more explicitly represent the signals you will be using," he says.

The decision tree-based methods APG uses to predict the future movement of the currency in a week or a month include random forests, which combine the output from multiple independent decision trees to arrive at the final output. Another technique it is uses is XGBoost, a fast approach to gradient-boosting decision trees—a method that includes an ensemble of trees, where each new tree corrects the error of the previous tree.

"You need to spend a lot of time defining well your predictors in a way that will help the machine learning methodology to find information rather than noise for making its predictions. Sometimes, as in the case of nowcasting, we would build a model that precedes the machine learning model to be able to apply it," Felix says.

Previously, the fund employed quantitative methods, such as rank and regression-based models, without using machine learning. The firm still uses these methods alongside machine learning. "Typically, the machine learning predictor enters the set of existing models that we have, as an additional predictor for the strategy. So it does not replace [other predictors], but it is used as a complement to the existing methodology. And we still believe that the previous methodology was OK. But investment strategies need to evolve. Marginally we are adding new predictors to this methodology, so it can keep producing returns. But it's not a matter of replacing it; it is a matter of actually making the strategy more robust," Felix says.

To arrive at a final score for its currency predictions, Felix says his team will "linearly combine the position sizes suggested from the machine learning-based predictor with the ones coming from our original methodology."

Using tree-based methodology has improved the prediction accuracy. "The machine learning approach for currencies has, on average, increased prediction accuracy by roughly 10% within our out-of-sample period, including its live period," he says. Out-of-sample is a measure of how accurately a model can make predictions for new data. wt

Investment bank wraps up major tech overhaul after Covid-19 setback

Stifel Europe weathered 2020 volatility and switched vendors in looking to simplify its middle- and back-office functions and increase tech investment. By Josephine Gallagher

Major technology overhauls like replacing middle- and back-office functions are complex beasts to tackle at the best of times, but trying to complete this kind of project in the midst of a pandemic comes with a host of new challenges. John Owen, COO at mid-market investment bank Stifel Financial, learned this the hard way. In March 2020, Stifel was just a month away from going live with the first stage of its implementation of Torstone Technology, consolidating its middle-office and post-trade operations in Europe, when the Covid-19 pandemic hit the continent.

From that point on, it was all hands on deck, Owen says, for the initial few weeks of the first lockdowns. The implementation had to be paused to allow the trading and risk desks to cope with extreme levels of volatility, and the workforce had to adjust to working remotely.

"The trading business went berserk. Volatility was all over the place, which was great for us, but our operations team had no time at all to even think about doing a parallel run or a conversion," Owen says.

Six months into the global health crisis, the operations teams adjusted to the new way of working and began revisiting the implementation. Owen says because the teams were dispersed, they had to become accustomed to running parallel tests remotely. This was done by what he describes as "over-communicating"—constantly being on open Zoom calls and having frequent check-ins at the beginning and end of the day.

The bank's implementation of Torstone Technology was interrupted by the pandemic

The project took an extra year to complete in Covid conditions, including all the necessary testing hurdles that had to be satisfied while working remotely. As a benchmark, a typical middle- or back-office migration for a small organization trading a simpler asset class like equities, can take anywhere between six and nine months to finalize, says Mack Gill, Torstone COO.

"This is the books and records of an organization," Gill says. "This is not a fancy tool in the front office or some risk tool; this is core. You have to make sure the testing is ready and make sure it works."

For Stifel, matters were made more challenging as its essential workers had to follow a rotation system for going into the office. Despite the implementation delay, Owen credits his team and Torstone for getting the project over the line. "Timing-wise, this is not something you could take as an example on the face of it. If you looked at it [in abstract], you would think it's incredibly bad, but actually, we've done an incredibly good job under the circumstances," Owen says.

As part of the first implementation, for the last six weeks, Stifel's fixed income and convertibles business has been live on Torstone's new middle-office solution, which automates allocation and confirmation processing, and its post-trade platform. In the next phase of the project, scheduled to be completed by September, the bank will migrate its Keefe, Bruyette, and Woods equities unit in Europe onto the third-party platforms, consolidating the various business lines and making them accessible via a single user interface.

Prior to this, Stifel used BNY Mellon subsidiary Pershing's outsourced solutions for its middle- and back-office functions. There were several reasons the bank decided to re-platform its European and UK businesses, including its need to scale, consolidate its product lines, and benefit from bespoke functionality. But Owen says flagging these needs became a challenge when competing for attention among a large group of clients.

"We were also one client of many. We happened to be one of the bigger ones, but Pershing as an organization is servicing all of its clients. So, anything that we needed was in a queue and our needs outgrew what Pershing could offer," Owen says.

Taking back control

Another driver for the bank to change providers was its need for better control over its technical footprint, its data, and its third-party vendors. Owen hopes that the bank's shift from Pershing's outsourced model to Torstone's cloud-based, software-as-a-service (SaaS) model will simplify its operations, enable the bank to have better control over its change management, and more easily switch out other third-party vendors, such as who it uses in the front office. Today, the bank uses Bloomberg and Thomson Reuters to trade and manage its fixed-income portfolios, Leversys for its convertibles business, and Fidessa for equities trading.

"The driver was our ability to be the master of our own destiny, rather than be in somebody else's hands, and everything stems from that. That is the eureka moment, and to be able to say, 'Look, we can do what we want in our own time,'" Owen says. wt

Show your workings: Lenders push to demystify AI models

Machine learning could help with loan decisions—but only if banks can explain how it works. And that's not easy.
By Steve Marlin

From derivatives pricing to credit card fraud detection—and a few places in between—artificial intelligence (AI) is extending its reach across the financial sector. But difficulties with explaining to regulators and senior management how self-learning algorithms work continue to hold back the use of machine learning in most banks' core business of lending.

"Credit underwriting is the highest-risk use of this technology and we would expect a great deal of explainability to be provided," says a model risk manager at a US regional bank. "We are beginning to use it for credit decisions but are very hesitant and are not yet comfortable that the benefits outweigh the drawbacks."

At the regional bank, the use of machine learning in credit underwriting has so far been limited to relatively small portfolios. The bank uses machine learning more extensively in other areas, including anti-money laundering and fraud prevention as well as marketing, where explainability is less of a concern.

"We do require an attempt at explainability from all machine learning models, but we acknowledge that different uses will have different needs in terms of the level of explainability," says the model risk manager.

An important part of model validation is to understand exactly why an algorithm produces a given result. A self-learning algo that spouts unpredictable outputs leaves the bank at a higher risk of losses.

In credit underwriting, where lenders assess the suitability of customers for loans, the stakes are even higher. Banks could leave themselves open to costly lawsuits if their models unwittingly discriminate against particular social groups.

But while the risks may be high, so too are the rewards. Machine learning could transform credit underwriting by helping banks automate much of the drudgery in assessing loan applications. Greater speed and accuracy in this area

could bring cost savings and a lower risk of loan losses.

So far, most applications of artificial intelligence in lending have involved decision tree methods, which use if/then parameters and are considered the most basic and transparent forms of machine learning.

JP Morgan has been developing gradient-boosted decision tree models—where multiple decision trees are combined to reduce prediction error—to generate proprietary credit scores for use in its consumer and community bank. The machine learning models are able to take in hundreds of attributes. The bank says the models provide a finer-grained ordering of risk than traditional credit scoring models that rely on logistic regressions.

The decision tree approach was specifically chosen because it is relatively easy to explain. "Explainability is a key consideration in choosing what tool to use for business problems. When the need for explainability is high you should choose tools accordingly," says David Heike, head of risk modeling for consumer and community banking at JP Morgan Chase.

Two of the most popular techniques for explainability are Shapley Additive Explanations (Shap) and Local Interpretable Model-agnostic Explanations (Lime). Lime is a statistical technique that analyzes a model's parameters to determine how it arrives at its outputs. Shap, which is drawn from game theory, explains the impact of each variable in a model on other variables, in the same manner as one would analyze the impact of one player in a sports team on other players.

These techniques do a good job of explaining the outputs of decision tree models. Things get much more difficult when it comes to deep learning—an advanced subset of machine learning that is starting to be deployed in other business lines, such as trading. Neural networks, one of the most common forms of deep learning, can find non-linear relationships in large swathes of data, making them potentially useful in credit modeling. But they are also opaque and difficult to understand. Neural networks can contain multiple hidden layers that transform the input data in ways that are difficult to trace, let alone explain.

Some firms are now searching for new techniques that can demystify the inner workings of neural networks so they can be applied to more sensitive tasks, such as credit decisioning.

Wells Fargo has developed methods for explaining a widely used form of deep neural network with rectified linear units (ReLU). One such method decomposes the network into an equivalent set of local linear models which are easier to interpret, the bank says.

If the technique proves successful, it may be an important step in helping lenders to explain complex machine learning algorithms in credit underwriting.

> "Fraud stands out as a particularly good use of machine learning given the dynamic nature of fraud attacks."
> David Heike, JP Morgan

Horses for courses

With regulators on both sides of the Atlantic scrutinizing the use of machine learning models, banks are adopting a "horses for courses" approach, reserving the more advanced techniques for less sensitive tasks with lower explainability requirements. The more sensitive the application, the easier it must be to explain.

Along the continuum of risk, banks tend to divide applications roughly into four broad levels. Level one applications—the riskiest—include credit scoring models. The use of poorly understood algorithms for consumer lending has the potential for serious harm to the lender as well as to the borrower. In the US, banks have a legal obligation to explain the decision to approve or decline an application for credit. The legislation is laid out in the Fair Credit Reporting Act and the Equal Credit Opportunity Act.

There are no equivalent laws in Europe, but fair credit is covered under the broader European Convention on Human Rights, with EU member states responsible for their own specific laws on discrimination.

Level two machine learning applications include fraud alerts and anti-money laundering systems, where the machine is making decisions that need to be acted upon in real time or near real time, and that could also affect customers.

Level three includes applications that could affect a firm financially but has no direct impact on customers, such as trading or internal stress testing.

The lowest risk category—level four—would include applications that have relatively little financial impact, such as product marketing. Here, banks are more free to apply more sophisticated but

> "We have zero appetite for black box models, and we do not buy or build black box AI models."
> Executive at a US bank

hard-to-explain machine learning techniques, such as artificial neural networks.

"For low-risk models, the techniques, the approach, the accuracy, the correctness of explainability is less demanding. For high risk, one should employ inherently interpretable machine learning models," says the head of model risk at a large US bank.

The first, and so far only, risk management application of neural networks within JP Morgan's consumer business is for signature verification. Neural networks are especially well-suited to finding patterns in unstructured data such as images, which contain thousands of seemingly random data points. JP Morgan uses a neural network algorithm in its consumer business to compare the signature on a cheque to past signatures to uncover inconsistency. If the algorithm is unable to determine whether the signature is genuine, it gets passed on to human operators for resolution.

"Fraud stands out as a particularly good use of machine learning given the dynamic nature of fraud attacks. Moreover, explainability is easier because you're explaining to sophisticated users who understand complex models," says Heike.

Although techniques exist for explaining the neural network's results, they are less well developed than for other types of machine learning applications. But because signature verification is a relatively low-risk application, it's not necessary to understand the intricacies of how it arrived at its decision.

Among the level three applications, banks are actively exploring the use of deep neural networks in areas such as derivative pricing, which has traditionally relied on a combination of classical models. These models include Black-Scholes, Monte Carlo simulations and finite difference method (FDM) techniques.

For derivatives pricing, Danske Bank uses a deep neural network that learns the pricing function from data. Once the pricing function is learned, it can be evaluated in near real time with different scenarios, orders of magnitude faster than Monte Carlo or FDM, helping to resolve computational bottlenecks. To some extent, pricing by machine learning is similar to traditional analytics, except that the pricing function is not derived by human mathematicians, but learned by machines from simulated data. The algorithm finds an intelligent manner to compute prices, but it does not determine or explain those prices in any way.

"Explainability is mostly a red herring, at least in this context," says Antoine Savine, chief quantitative analyst at Danske Bank. "Prices are explained by the pricing model that simulates data, not the algorithm that learns the pricing function from that data. In this context, machine learning is just a way to compute prices efficiently, like FDM or Monte Carlo, and like with numerical methods, the key notion is not explainability but convergence and error analysis."

Savine adds that explainability may be critical in other applications of machine learning in finance, such as trading strategies, credit ratings or synthetic data generation.

Cutting through the noise

For the most risk-sensitive applications, such as lending, banks will choose techniques that are the easiest to explain using post hoc techniques or techniques that are inherently self-explanatory.

Traditional consumer credit underwriting models, such as Fico scores, rely on statistical techniques such as logistic regression. Here, the explanation can be extracted directly from the model by measuring the sensitivity of the output to changes in inputs, such as income or amount of debt outstanding.

Explainability techniques in machine learning essentially freeze the model and take it apart in the manner a mechanic would a car. The problem is they often have a hard time differentiating between what's substantial and significant from what's affected by noise. When applied to financial data, which is inherently noisy, it becomes doubly difficult.

The situation is analogous to asking Google to provide driving directions from New York to Chicago. After setting out, an accident snarls traffic on the chosen route, at which point Google comes out with a whole new set of directions. With route-planner algorithms, the user knows that the new route is the result of a traffic accident. But with machine learning models in finance, the cause of a new output is not so easy to determine.

"When you're deducing the behavior of the model, and you train the model again you get a completely different set of explanations. How can you reconcile very different techniques all pointing at different targets and none of them getting at the core issue, which is noise and whether you have enough observations for that particular explanation to be statistically significant?" says Matthew Dixon, a professor of computer science at Illinois Institute of Technology.

In addition to the riskiness of the application, the choice of explainability technique hinges on whether the explanation needs to be global or local. Global describes how the model behaves under all possible assumptions, and local describes how the model arrived at a particular decision.

For example, if a bank wanted to measure the sensitivity of a consumer lending model's outputs to factors such as disposable income or GDP, then a global technique is required. On the other hand, if the bank wanted to understand how changing a variable would impact a particular lending decision, then a local technique is needed.

Shap and Lime are the most popular techniques for local explainability.

An example of a global explainability technique is "partial dependence plots," which measure the sensitivity of a model's output to a particular set of inputs. Other techniques exist, including relevance, sensitivity and neural activity analysis, as well as fitting a simple model on top of a more complex model. There are also visualization techniques for explainability that are applicable to decision tree-based models.

Part of the problem is that each explainability technique tends to come up with a different explanation for a model's output. Therefore, it is common practice to use more than one technique. "All of these post hoc explainability approaches are approximations and may encounter computational problems. That's why they often don't agree with one other. So one needs to be careful and should apply multiple techniques," says the head of model risk at the large US bank.

Wells Fargo has developed a technique that pares back the number of linear equations that a deep neural network produces, helping users interpret the results of the model. Simply put, a neural network attempts to mimic the activity of neurons in the human brain. Each neuron—or node—receives an input and performs a calculation to decide whether that input meets a predefined threshold. If the threshold is met, the node "fires" and produces an output that travels to the next node. The Wells Fargo model uses the ReLU mode of transmission—a widely used activation function for neural networks.

What Wells Fargo has done is to trace the journey that the information takes through each layer of nodes, and convert the data into linear equations—thousands or even millions of them, depending on the size of the net. But because many nodes have made the same decision, many of these linear equations are the same. So the equations can be sorted into groups, in a process known as regularization.

The sorting process enables the bank to reduce the number of equations and nodes—but without affecting the performance of the network. With a smaller number of linear equations, the structure of the network is more interpretable. At least, that's the theory.

Wells Fargo has since applied the technique to convolutional neural networks for natural language processing.

In determining where and how to apply machine learning techniques, modeling teams work closely with business lines to ensure the model risk management team has access to developmental code and a clear description of model features. Most banks will not buy any externally sourced models that do not meet these qualifications. "We have zero appetite for black box models, and we do not buy or build black box AI models," says an executive at a third US bank.

The policy of avoiding black box models proved to be prescient during the Covid-19 pandemic, when AI-based fraud detection systems were unhinged by the changing patterns of customer behavior stemming from the lockdowns. Typically, when a person uses a debit or credit card, whether in person or online, a model predicts in less than one second whether the transaction is fraudulent, and will freeze or flag the transaction. Through testing, the bank determined that the most predictive variable for fraud pre-Covid was whether the transaction was card present or card not present.

Realizing that its fraud detection models would lose their predictive power, the bank bought vendor transaction data covering states that were affected by stay at home orders, and by mid-April, after the models had been recalibrated using that data, it became apparent that card present and card not present was no longer one of the most predictive variables. "We performed manual and automated reviews to assess every variable and every AI model at the bank," says the executive. wt

The market data vending machine: The pros and cons of self-service procurement

Brokers and exchanges have begun rolling out 'self-service' portals that allow clients to choose data and services on an a la carte basis. Opinions vary on whether they are the Holy Grail or a poisoned chalice. By Max Bowie

For market data management professionals, it's the Holy Grail: being able to buy the data you want, when you want it, and pay for what you use, rather than paying a fortune for always-on, voluminous datafeeds or bundled services of which a firm uses a fraction.

While pay-as-you-go pricing is not yet a reality, some data providers—notably, among brokers and exchanges—are laying the foundation by providing "self-service" portals where clients can subscribe to content and services without having to contact a sales rep, streamlining previously manual permissioning and onboarding processes.

Industry observers say this approach represents a shift away from old-style commercial agreements, and is more geared toward clients' needs, creating greater flexibility for consumers.

For example, in April, inter-dealer broker Tullett Prebon combined its data and analytics, post-trade services, and investment analytics offerings (acquired from Liquidnet) under a new brand, Parameta Solutions. Part of the reasoning for this was that data is at the heart of many of the broker's non-data offerings, which in turn feed information into the data business.

Along with that, the broker rolled out a self-service portal for the unified services, allowing clients to choose from its services online, and providing a choice of delivery mechanism as firms increasingly move to the cloud. The offering is not intended to replace existing ordering and delivery methods, but to adapt to clients' changing needs.

"Clients can still access our data via our data partners, but we also want to meet new requirements and be able to deliver data to clients where they are," whether on-premise, in offsite datacenters, or in the cloud, says Roland Anderson, CIO of Parameta. "It's all about efficiency and flexibility as companies develop their technology stack."

The portal allows users to browse the broker's content, obtain samples of datasets for testing, and execute a contract online. Parameta then delivers the data directly into firms' Amazon Web Services S3 cloud storage. Anderson says that from conception, Parameta envisaged the portal as being cloud-native.

"We're producing data in the cloud, and we wanted to make it available in the cloud," he says. "We're trying to recreate an Amazon-style shopping experience."

The project was something of a technical departure for the broker. To create the portal, Parameta set up a new internal team to work on the user interface and enlisted a group of "external data experts" with experience in building e-commerce platforms for financial firms. But it also employed technologies more associated with Facebook and Airbnb than capital markets, including the analytics query language Presto, Airflow for reporting, and the Apache Kafka open-source event and data streaming platform.

Parameta isn't the only broker data arm making use of self-service portals.

> "We're empowering the customer, giving them access to pricing and educational tools—from the basics, like learning what a call option is, to more sophisticated options strategies. And with execution, which we recently rolled out, we're giving customers the ability to place their own executions, and choose strategies, from conservative to aggressive."
> **Karrissa Allyn, StoneX Markets**

Roland Anderson
Parameta

StoneX Group, formed by the merger of Intl FCStone and Gain Capital, recently rolled out its MyStoneX self-service portal, aimed at the firm's commodities trading clients. The portal provides a single point for clients to access "market intelligence, risk measurement, scenario analysis, and market execution" for futures and options trading.

All the firm's over-the-counter (OTC) clients receive access to proprietary content on the portal at no extra charge. To place orders through the platform, traders place a request, then must submit to a compliance and risk process. But in keeping with the aim of streamlining access, "It's a quick turnaround," says Karrissa Allyn, vice president of business strategy at StoneX Markets.

"We're empowering the customer, giving them access to pricing and educational tools—from the basics, like learning what a call option is, to more

sophisticated options strategies," Allyn says. "And with execution, which we recently rolled out, we're giving customers the ability to place their own executions, and choose strategies, from conservative to aggressive."

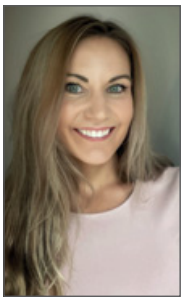
The tools available via MyStoneX allow its brokers to have more meaningful conversations with clients about how to customize trading strategies and enable the firm's clients to more easily track commodities and prices, provide a better service to their customers among grain and other commodity producers, and ultimately to win more business and trade greater volumes.

"This was a way for us to do something new and stand out against our competitors. The biggest thing is the ability to get that information quickly because when a producer calls, they want to make decisions quickly," says Matt Zeman, vice president of grain at Iowa-based Cooperative Farmers Elevator, which uses the StoneX platform.

Though MyStoneX is new, it's been in the works for around four years and is part of a larger digitization project at the broker. "We're in the process of digitizing our entire product line—from vanilla swaps through to more complex products. I would say we're now 70% along that process," says Mark Mauer, CEO and president of StoneX Markets, who says the portal will establish StoneX at the forefront of a commodities industry that is often slow to adopt efficiencies.

Part of that timeframe was spent reassuring the firm's brokers—who were initially skeptical—that the platform was not intended to remove brokers from the process, but rather to free them up and give them more time to focus on relationship-building with clients. "This was driven by demand from our customers and brokers. After the initial build, our brokers realized the value this could add to their workflow, and so following that, enthusiasm for more features and functionality grew exponentially," Mauer says.

But self-service portals aren't totally new: CME Group has been operating its self-service data portal for five years, though it was able to get a head-start over others by employing third parties—and, like Parameta, making extensive use of the cloud—to provide specific components, rather than building everything in-house.



Karrissa Allyn
StoneX Markets

Mark Mauer
StoneX Markets

Bernardo Santiago
S4 Market Data

CME's initiative began as a more efficient way of distributing data, but now allows the exchange more time to focus on product development than on sales and distribution.

Traditionally, to obtain historical data from CME, a client would need to phone a salesperson to license and buy the data. Then the exchange would load data onto a hard drive and physically ship it to the client.

"A few years ago, we set out to figure out what's in the best interests, not only of CME, but of our clients, and how could we deliver services quickly and easily to clients," says Trey Berre, global head of CME Data Services at CME Group.

In 2016, the exchange enlisted Canadian data technology vendor TickSmith to revamp DataMine and provide a more modern and—key to its expansion plans—scalable platform. TickSmith deployed its Enterprise Data Web Store and its Gold (Gather-Organize-Leverage-Distribute) sales platform. Now, existing clients can get their hands on data within minutes, while new clients can get access within a couple of days, after completing CME's know-your-customer (KYC) requirements.

The following year, CME tapped DataBP, which usually builds marketplaces for organizations to automate data sales activity, to build a web-based data licensing portal to automate the manual processes of signing and returning contract documents, drastically reducing the time it takes for clients to license data. And in 2018, it added the ability to pay for data using a credit card.

"For large banks, their data requirements are complex and take more time. But for smaller institutions, this has been very useful. And if they want to add other products, they just need to go through a maintenance procedure that can be completed in hours, if not minutes," Berre says. "Making data available on a self-service basis helps us focus more on building better products for clients. So we feel it's a win for everyone."

Since then, CME has used DataMine to expand the range of datasets it offers via self service, including 28 proprietary datasets—including top-of-book and market depth data, time and sales data, US Treasury data from Brokertec, as well as liquidity tools—and data from 17 partner organizations ranging from exchanges to alternative data vendors. The exchange also offers a free 30-day trial of the datasets available.

Berre says one of the aims was to provide more variety in how clients could access CME data, and also to appeal to a wider variety of potential clients. The exchange initially envisaged that the portal would appeal to sophisticated retail traders. And while it has seen some takeup from that market, it has also seen demand from crypto trading firms wanting futures and options data—not just on bitcoin futures, for example, but also on gold futures and CME's Dow Jones Indexes products because of the interplay between these markets—as well as from firms wanting to cut costs by leveraging cloud delivery to reduce their legacy infrastructure footprint.

"We are definitely seeing more of these data marketplaces, but I do feel we are a way off from seeing the true impact," says Bernardo Santiago, founder of market data consultancy S4 Market Data. "These self-service portals are seeking to add a level of ease to the data acquisition process," which is evident for the data providers, though Santiago says he is yet to see that translate to ease for consumers, as the data onboarding process must still be managed by consumers outside of the vendor's self-service portal.

Cautionary Notes

While data professionals say the portal-based approach offers potential benefits in terms of efficiencies and time savings, they note that it also raises some red flags from a data management perspective. First, depending on who has access to the portals, firms could see unchecked rises in spend on data—for example, if end users on trading desks or data science teams are allowed to sign up for content without adhering to established processes governing how a firm buys data.

"If you open this up to end users, it potentially conflicts with the vision of the market data group around

"Self-service portals do make life easier because they remove the need for phone calls and emails, so it's better for those selling data because they can automate the sales process. But it's not necessarily better for consumers, because while they do want automation, they also want straight-through processing. And with these portals, you still need to have humans involved. Clients still need to re-key everything that they would have called their rep about in the past." **Richard Mundell, TRG Screen**

controlling costs. If business users want to buy data to feed into applications, the market data group could get blindsided by costs and by where data is being used, and possibly get tripped up during an audit process," says one market data manager.

Ensuring that the process is properly set up contractually and can be accounted for in a firm's inventory management system will be key, the market data manager adds. "If you only put it into the hands of market data folks, they should be privy to the terms and conditions of any given agreement," he says.

Typically, end users request data via a central procurement or market data management team, which ensures users are properly entitled and licensed, that the contract is managed properly, and usage and costs are tracked via the firm's inventory management system. This process, in particular, would at best be manual when acquiring data via self-service portals, requiring data managers to update their inventory systems, and at worst might end up being inaccurate if end users buy data without enlisting market data management and properly recording spend and usage.

"In the software-as-a-service (SaaS) world, you hear about 'shadow IT' where people spend money on technology themselves. It always happens, and market data teams would have to scoop that up and manage it centrally. Plus, end users have been well trained to not do that. We don't see these portals as ways for vendors to work around data teams, but rather that they realize that ordering data in person or by phone is not the most efficient way to work anymore," says Richard Mundell, chief product officer at inventory management software vendor TRG Screen.

Mundell draws another comparison to the SaaS world, noting that providers ranging from Microsoft to Salesforce to cloud operators allow clients to provision access to services using API calls, whereas client portals remain a manual link in the purchasing chain.

"Self-service portals do make life easier because they remove the need for phone calls and emails, so it's better for those selling data because they can automate the sales process. But it's not necessarily better for consumers, because while they do want automation, they also want straight-through processing. And with these portals, you still need to have humans involved. Clients still need to re-key everything that they would have called their rep about in the past," Mundell says.

Richard Mundell
TRG Screen

Trey Berre
CME Group

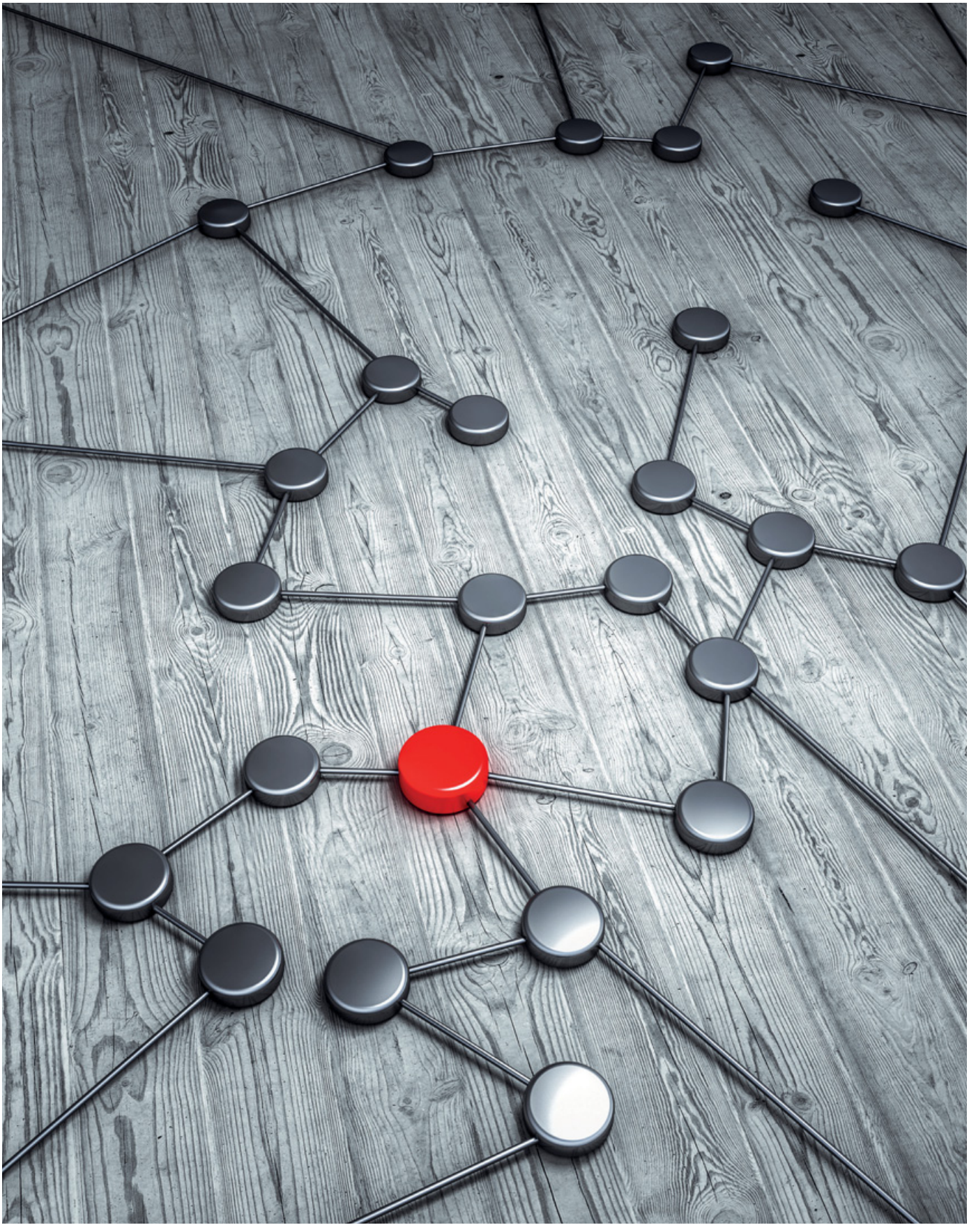
For example, if a bank uses TRG's Optimize Spend platform, a business user can log into the platform, see what services the firm has available, and request any service available in its current catalog of services. That request is routed for approval, and to the market data team to acquire the data. But that's where the process hits a manual roadblock.

"When it reaches the market data team, they still have to go to these portals or phone or email the vendor. So the fulfillment remains manual. That's the piece that could be automated," Mundell says. "But the vendors don't have APIs set up to handle those requests. The thing is, all these portals have APIs on the back end—they just don't expose them to customers."

In addition, Mundell says that by not leveraging APIs, data providers are potentially "leaving money on the table," because APIs would allow them to raise their profile and expose their services to potential users who may be unaware of their existence. "If vendors provided APIs, in addition to automating fulfillment, it would allow end users to browse not only what their firm has already licensed, but the entire catalog of all services available from all vendors who expose their APIs," he says.

But if brokers and exchanges can allay firms' data management concerns, the market data manager says firms should test the portals. "You would definitely need to trial it. You would need technical people, representatives from the user community, finance and admin people, and information security staff, and bring them all to the table because no single person will be able to cover all the bases and anticipate any potential pitfalls," he says.

At the end of the day, self-service portals aren't the Holy Grail of market data provision. Not yet, anyway. But they're certainly a step toward making that a reality. wt



As fixed income edges toward automation, the interop movement is cutting in

Valantic FSA, a European solutions provider, wants to remake the fixed-income tech scene in interoperability's image, taking on incumbents like Ion Group. By Rebecca Natale

Though interoperability has become the norm in equities trading technology, it has been slower to make headway in fixed-income markets. While a higher percentage of bond trades are executed electronically these days—due in large part to the Covid-19 pandemic and remote work setups—technology providers will go to exhaustive lengths to develop new fixed-income platforms, touting machine learning, innovative protocols, and other unique capabilities to entice users.

Whether niche vendors and new entrants—such as LTX, operated by Broadridge Financial Solutions; the Project Octopus consortium, a bank-led platform for trading collateralized loan obligations; Trumid, which secured a $50 million funding round earlier this year; Broadway Technology, a newly independent company following a partially failed takeover by Ion Group; or Ion Group itself, a giant in fixed-income technology that has grown through numerous acquisitions—can push the bespoke, fragmented world of fixed-income trading fully into the digital age, remains to be seen. Likewise, industry stalwarts Bloomberg, Tradeweb, and MarketAxess have their own horses in the race.

Despite efforts to "electronify" workflows, most fixed-income traders seem to choose their old-school phones and messaging apps over disruptive products to make their trades. Perhaps then, some industry participants say, existing platforms and technologies can work smarter, not harder, by allowing integration with one another, as opposed to searching for the elusive unicorn.

If you can't beat 'em, join 'em

Meet Valantic FSA, the financial services automation arm of European consulting and solutions company Valantic, which, over the last 18 months, has been expanding its customer base outside its primary region of Switzerland, Germany, and Austria. The unit, which combines fixed-income trading components with mesh technology to deliver solutions via HTML, offers two main automation platforms—one for front-office trading across asset classes, and another for post-trade functions—and low-code tools for building custom workflows.

Throughout its continued expansion—the FSA group counts 100 clients in 15 countries—the company has decided to place all bets on "choice" as its operating philosophy, says Joachim Lauterbach, CEO of Valantic FSA. For him, choice means two things: fewer and less demanding terms and conditions that lead to lock-in, and interoperability, including not just open APIs and the desktop application breed supported by the Financial Desktop Connectivity and Collaboration Consortium (FDC3), but pre-built modules and frameworks, which clients can stitch together to create their own user experience.

"[We are] deploying IP where a financial institution might not have the option of building it themselves, and combining it with the ability to interoperate with other systems. And on top of it is a framework that allows them to build something new—build their own applications—on top of the data mesh we provide," Lauterbach says.

Clients have another choice. Because Valantic has partnered with some of the major desktop application interop vendors, including OpenFin, Glue42, and Cosaic's Finsemble, clients can plug in either of Valantic's standardized platforms to their interop vendor's environment, or use Valantic's framework to build their own HTML5 front-ends, which can also be supported by these vendors' containers. In other words, the front-end interop is already a given. The back end is at the core of Valantic's pitch to the industry.

Steve Grob, founder of fintech consultancy Vision57, has been working with Valantic FSA for the past year, as part of a handful of his wider industry interoperability projects such as with Glue42 on legacy order management systems, an example of server-side, or back-end, interop. The server-side interop play aims to build off of what the web- and browser-based brand of interop has accomplished—getting

"The Cosaic–Glue42–OpenFin world is clearly in this world of desktop interop. And that is the beginning of something even bigger, I believe, that's going to be about data interop. And the idea is you've got these vertical stovepipes of tech, typically by asset class, or by business line, or what have you. And what you want to be able to do is build and recreate different business units out of them in order to do the right job for customers, but you can't because the data is stuck in system A, and it won't get into system B or system D. The magic is when you separate the data from the applications, suddenly you can re-combine that data in any way you want."
Steve Grob,Vision57

legacy, monolithic platforms like order management systems and terminals to talk to one another at the discretion of the user, not the operator.

"The Cosaic–Glue42–OpenFin world is clearly in this world of desktop interop. And that is the beginning of something even bigger, I believe, that's going to be about data interop," Grob says. "And the idea is you've got these vertical stovepipes of tech, typically by asset class, or by business line, or what have you. And what you want to be able to do is build and recreate different business units out of them in order to do the right job for customers, but you can't because the data is stuck in system A, and it won't get into system B or system D. The magic is when you separate the data from the applications, suddenly you can re-combine that data in any way you want."

The result, Valantic bets, is a vast reduction in the technical debt that financial firms manage. Indeed, Valantic has had to make bilateral agreements with other providers, including Bloomberg, Fidessa (now owned by Ion Group), FlexTrade, Overbond, and others, to enable some of this data interoperability, but the main intent of Valantic's platform is another variation on choice—not simply choosing parts, but choosing how those pieces are used once they're brought into one's own organization.

Valantic surely has lofty aspirations, and Lauterbach is aware he runs the risk of positioning his company as being all things to all people. But, he contends, that is simply where the market's needs have pushed them.

"It wasn't the case that we were sitting in front of a white sheet of paper, draw-

ing things, and said, 'So how can we be Mr. Superhero or Mrs. Superhero for solving every single problem that's out there?' It is more of an evolution for us," he says.

Elephant in the room

It would be remiss to discuss the fixed-income trading software scene without acknowledging the elephant in the room: Ion Group, the biggest vendor in the space by a long shot, and the target of much ire from users. Its MarketView product is said by one rival to be in use at 75% of the world's top 50 sell-side banks, and dealers, which from smallest to largest are estimated to spend anywhere from $2 million to $12 million per year licensing Ion's software.

As well as trader screens, pricing, and risk analytics, Ion packages also include the Ion messaging bus—used to ship data and orders between different systems—and connectivity to scores of trading venues around the world. It's this offering that Tommaso Di Grazia, head of fixed-income product development at Ion Markets, says allows the interoperability and choice that Lauterbach and Grob say are lacking in the market.

Steve Grob
Vision57

Di Grazia says that on the front end, there's the Ion desktop, where APIs unite the many applications Ion has acquired over the years. Users can glue and stitch modules together as well, so that if customers change the theme of the UI, for example, in one app, all others will follow suit. If a user wants to source data from their own repository, they can join it with Ion's data without going to the back end, using the Ion bus, a middleware component through which users can create C# plug-ins and represent corresponding data in the UI.

Di Grazia says that when it comes to interoperability with third-party systems and tools, Ion's allowance of openness is more hands-off.

"So far, we've seen customers taking [on] that piece of work. In general, the expertise is built in-house and not outsourced and not offered to a third party," he says. "If you open up [the desktop] of a large client, there's a huge amount of custom components. We don't even know what they're doing."

The path to interoperability is nearly seamless now in the mainstream commercial tech industry, as the average smartphone can contain an individual's entire life. Financial technology's pursuit of it has paved bumpier roads, even in equities, where it's more established. In fixed income, modern technology may still take years to entrench itself, or it may never change—because, other than a pandemic, not much else has worked so far. wt

Can interop connect the bond market better than consortiums? (Yes)

We know that fixed income is fragmented and complex from both workflow and market structure perspectives, especially when compared to equities and foreign exchange. This has led to the launch of numerous trading platforms and venues, as well as data offerings that aim to help automate the space and thus improve liquidity and drive down margins.

For example, a few months ago, *WatersTechnology* broke the news about a project between Citi, Bank of America and others to build a new execution platform for fixed-income markets, initially focusing on collateralized loan obligations and syndicated loans. If this consortium of major fixed-income players can pull this off, it could prove revolutionary and disruptive. But, as we've written about before, consortiums often fall short or get spun off into independent platforms that eventually raise fees—thus angering the banks—and so the banks start thinking about coming together to build a new consortium-led platform. It's a flat circle.

But what if big-bang projects aren't the answer? What if the data and technology to improve electronification and liquidity already exist, but it's simply a matter of connecting disparate systems? What if simple workflow interoperability is the answer to the woes of the fixed-income market, rather than new platforms, venues, and tools? As Reb Natale writes in the article accompanying this column, what if we got technology to work smarter, not harder, by allowing integration with one another, as opposed to searching for an elusive unicorn?

"The Cosaic–Glue42–OpenFin world is clearly in this world of desktop interop. And that is the beginning of something even bigger, I believe, that's going to be about data interop," says Steve Grob, founder of fintech consultancy Vision57, and someone who is a big believer in interop. "And the idea is you've got these vertical stovepipes of tech, typically by asset class, or by business line, or what have you. And what you want to be able to do is build and recreate different business units out of them in order to do the right job for customers, but you can't because the data is stuck in system A, and it won't get into system B or system D. The magic is when you separate the data from the applications, suddenly you can re-combine that data in any way you want."

As Reb explains, there's a major barrier to interop in fixed income: The biggest trading platform provider, Ion Group, has been acquiring vendors in the space seemingly every month and there's a perception in the industry that the company is not interested in playing nicely with others. So what happens next? A consortium project!

Tommaso Di Grazia, head of fixed-income product development at Ion Markets, told Reb that while they use APIs to connect systems, when it comes to interoperability, Ion allows this to happen, but the vendor is also not actively pushing these projects—it has to be client-led.

"So far, we've seen customers taking [on] that piece of work. In general, the expertise is built in-house and not outsourced and not offered to a third party," he says. "If you open up [the desktop] of a large client, there's a huge amount of custom components. We don't even know what they're doing."

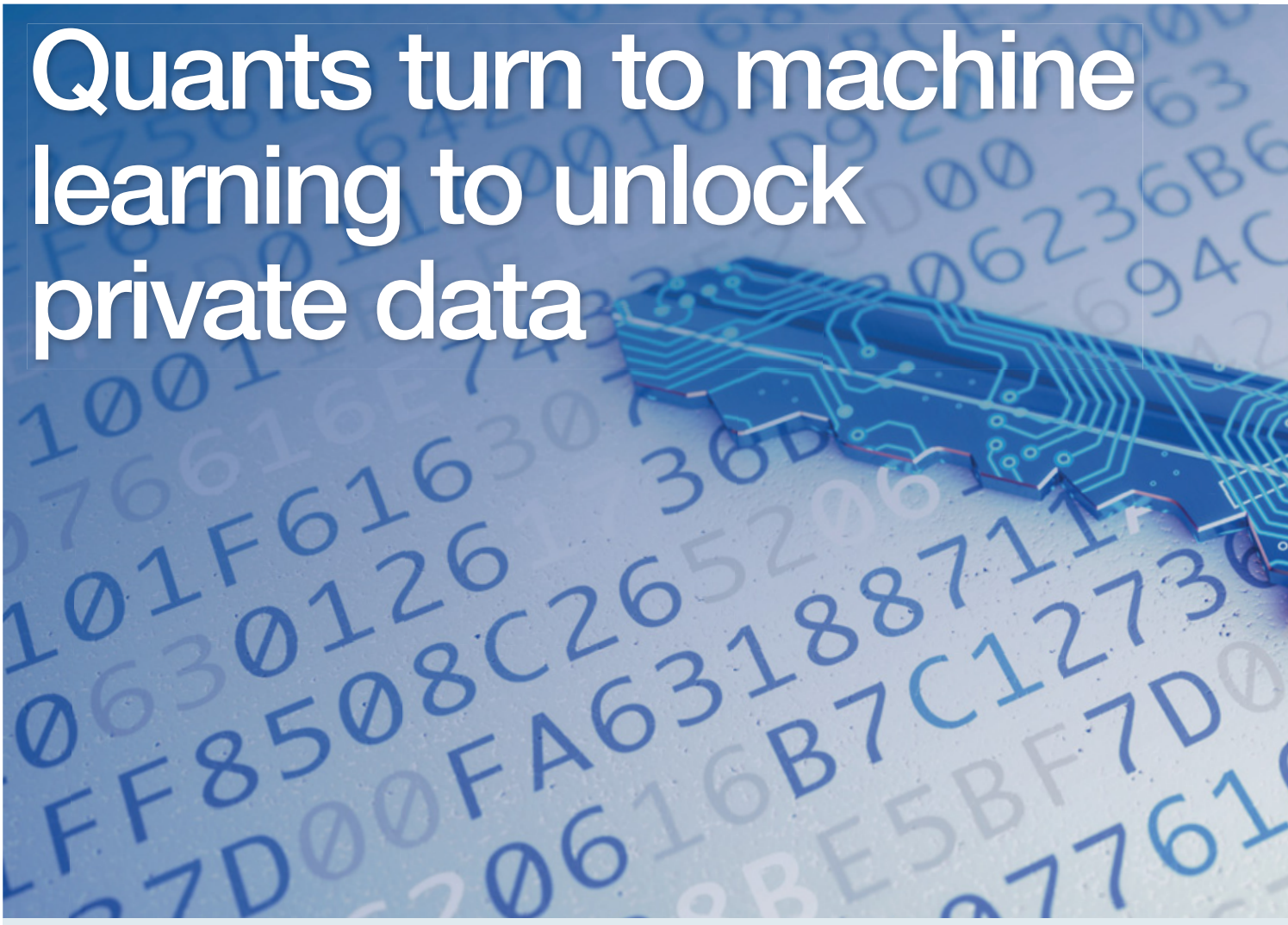
To be fair, Ion gets painted as the big bad wolf—and there are certainly reasons for that—but they are not alone in the fixed-income space when it comes to being closed off. Bloomberg is a giant as well, and while it has slowly opened itself up a bit to interop, its bread and butter Terminal is still a mostly closed-off ecosystem. Additionally, there's still more incentive for the likes of Broadridge, MarketAxess and Tradeweb to protect their respective turfs. And even a company like Broadway Technology, which has been essentially reborn after an acquisition attempt by Ion, has been more closed off with its Toc platform, but it looks like that might be changing with this new iteration of the firm.

There's good business in platform creation. It's exciting to see how new tools can help simplify fixed-income workflows. Consortiums can be energizing for the tech teams involved. But at the end of the day, I feel like we can't see the forest for the trees.

More and more data is coming to market, and there's rapidly growing hunger for new sources of it. New forms of AI are proliferating the market, and the use of open-source tools is helping to democratize—to a certain extent—machine learning and its cousin, deep learning. And at the same time, trading firms want to diversify their trading strategies and the regions they play in.

To me, interoperability is how you capitalize on those trends—not ambitious new platforms and venues. Think I'm off base? Let me know: anthony.malakian@infopro-digital.com.

Quants turn to machine learning to unlock private data



Replication could allow financial firms to use—and monetize—data that was previously off-limits. By Rob Mannix

When an investment firm wanted to find out how a new breakfast menu at Wendy's might affect the fast-food chain's bottom line, it looked for the answer in time-stamped credit card transaction data.

The data was anonymized, of course. Credit card companies remove sensitive information and add statistical "noise" to this type of data before selling it to investors or even sharing it internally. But these anonymization techniques are not foolproof, and nervousness about privacy breaches has held back the use of transaction data in areas such as investment analysis, fraud detection and the development of execution algorithms.

A new idea could change that. Rather than anonymizing datasets, financial firms are looking at replicating them. Machine learning algorithms can synthesize new, artificial datasets that are completely different from the original, while retaining the same statistical characteristics. Because the new data is essentially fake, it can be shared at will.

"The point about synthetic data is that it can remove sensitivities around personal information but preserve the signal," says Harry Keen, founder of Hazy, a UK-based synthetic data firm that works with financial services firms.

The approach could make it easier for investment firms to develop strategies using data that was previously off-limits. It could also allow them to test execution algorithms and other third-party services on their own data before signing expensive contracts.

Banks could use the technique to monetize proprietary data from their retail or trading units and to forge partnerships with fintechs.

Fidelity International is already testing the technology. "There's appreciation at senior level that this is the way forward," says Erik Mostenicky, a senior associate in the firm's strategic ventures group, which invests in businesses that are of strategic importance in asset management. A budget has been agreed and a team is working on a proof-of-concept. The plan is to put data anonymization using synthetic data "into production" by the end of the year.

Other large financial firms are making similar moves. Data scientists from American Express described how synthetic data can be used in risk management at the NeurIPS conference, an

annual gathering of machine learning experts, in Vancouver in 2019.

A team from JP Morgan gave a presentation on ways to synthesize data from order books and customer transactions at the same conference. "Financial services generate a huge volume of data that is extremely complex and varied," the bank's data scientists wrote in a subsequent paper on their research. "Data sharing within different lines of business as well as outside of the organization is severely limited." They called for more research into ways of synthesizing financial data to overcome these limitations.

Quants at Standard Chartered have also written and spoken about using synthetic data to anonymize sensitive information.

> "There's appreciation at senior level that this is the way forward." Erik Mostenicky, Fidelity International Strategic Ventures

Firms are cagey about what they're doing—American Express, JP Morgan and Standard Chartered declined to comment for this article—but anecdotal evidence of interest in fake-data anonymization on both the sell side and buy side is growing.

Fernando Lucini, head of data science at Accenture, says his team is fielding five or six enquiries a month about synthetic data. A year ago, they had none.

Mostenicky is convinced other financial firms are trying out the technology, even if they won't say so publicly. "I'm 100% sure it's happening," he says.

Behind the mask

The standard approach to anonymization involves stripping datasets of fields that could be used to identify individuals. More sophisticated methods add statistical "noise" to data to render any single record meaningless while ensuring the composite retains its value.

But these techniques have their limits. Well-known instances of so-called de-anonymization, where hackers have been able to pick out individuals from supposedly anonymized data, include tracking of celebrities' taxi trips and identification of the authors of Netflix film recommendations. Research has

> "With synthetic data, you can share samples without needing to sign non-disclosure agreements." Gautier Marti, Abu Dhabi Investment Authority

shown that in 90% of cases, the date and location of four credit card transactions is enough to reidentify an anonymized cardholder.

New laws such as the EU's General Data Protection Regulation and California's Consumer Privacy Act impose stiff penalties for such failures.

Legacy anonymization methods have another problem: They can mask the data too much. "You lose the pattern you need to validate investment strategies," says Mostenicky. In some cases, a majority of data points might need to be removed to protect privacy.

Mostenicky illustrates the point with a trivial example: After masking, the statement "BMW, Mercedes and Audi are German carmakers with billions in revenue," might become, "A, B and C are German objects of size."

Even sophisticated methods of anonymization such as homomorphic encryption come up short in finance, he says. Homomorphic encryption is a technology that encodes data in a way that cannot be de-anonymized while still allowing for calculations to be performed on it. The downside of this method is that it increases the size of the dataset, making computations thousands of times slower.

Anonymizing with synthetic data, by contrast, gets away from sharing the original data entirely.

"Our software takes the original training set as learning material," explains Alexandra Ebert, chief trust officer at Mostly AI, a fintech that has developed anonymization software for financial services firms. "Our deep learning algorithm identifies the patterns, the correlations, and understands how the customers behave, and what's logical for them. Once the training process is finished, a completely separate synthetic dataset is generated from scratch that has the same characteristics."

No single record in the synthetic dataset matches an original, but the synthetic dataset is still as useful as the original for training machine learning algorithms or for analytics, claims Ebert.

There are several machine learning techniques for creating artificial datasets. Standard Chartered prefers so-called Boltzmann Machines. Some practitioners advocate using variational autoencoders. An advanced technology—generative adversarial networks—is the same method used to generate viral deep-fake TikTok videos of Tom Cruise.

Opening doors

For investors, synthetic data opens the door to testing datasets more easily.

"When a fund wants to trial alternative data, usually it takes time," says Gautier Marti, a quant researcher and developer at the Abu Dhabi Investment Authority, and an expert in ways to replicate complex datasets. "With synthetic data, you can share samples without needing to sign non-disclosure agreements and so on."

Marti has looked at different ways that synthetic data might be used in finance and sees anonymization as the most obvious application.

Hedge funds making big investments in alternative data want assurances that the supply of data won't be cut off because customers change their privacy settings, says Lorn Davis, vice-president of corporate and product strategy at Facteus, a company that anonymizes card transaction data.

Investors also face restrictions on the use of existing data that synthetic data might help overcome. "Data from third-party providers like Bloomberg, Markit and others comes with huge restrictions," says Mostenicky. Because of legal, compliance and operational hurdles set by vendors or their own organizations, data "owners" within buy-side firms are "super anxious" about using data in new ways, he says.

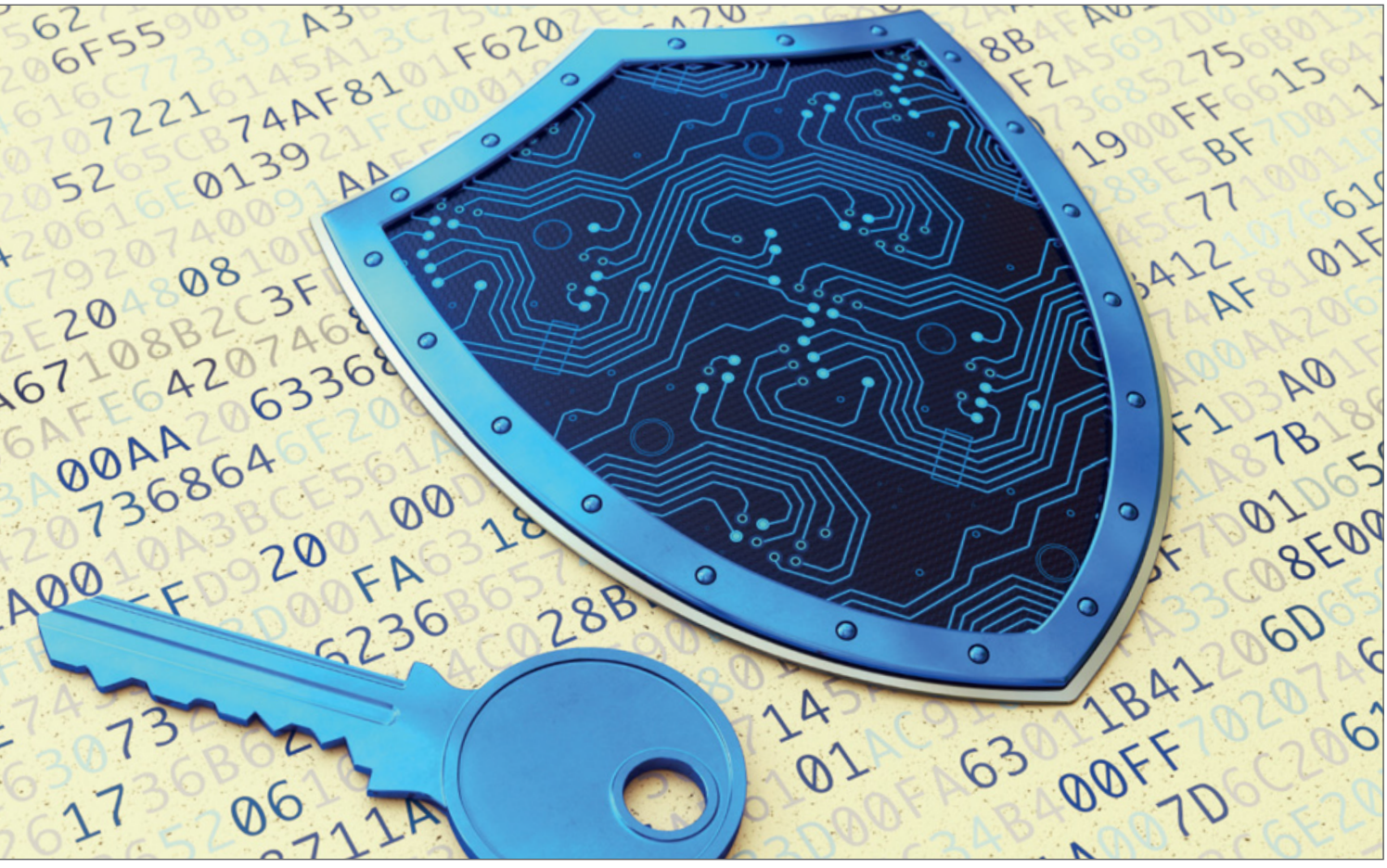
Using data to verify whether a startup seeking investment can deliver on its promises, for example, could be classified as commercial use, which is often prohibited under contracts with vendors.

"With synthesizing, you can replicate the shape and form and size of the data that you're trying to use without the hassle of having to ask a vendor like Bloomberg as well as your legal team whether you can use it," Mostenicky says.

Investors could also use the technology to share their own data with vendors and external service providers. "A firm might say its technology will generate signals that tell you when your traders need to sell and when to buy," Mostenicky says. "To test that you want to provide proprietary datasets on the performance of your investment funds. But you want to make sure you don't disclose everything.

"With synthetic data, you can allow the external firm to use your synthetic securities and investment holding data and provide signals on the data they see."

In the most advanced cases this could extend to creating a "sandbox" environment where external parties gain access to a range of anonymized data such as investment data or pricing data for use in proof-of-concept exercises.



"You can give them access to the data without revealing any of your investment strategies or your allocations," Mostenicky says. "That can shorten the time to determine whether a startup or possible partner is relevant or not."

For banks, too, a fail-safe method of anonymization would grease the wheels of collaboration with outside partners, specifically fintech companies sweeping the industry with offerings such as fraud detection systems or loan-default prediction models. Often those companies employ machine learning and so need access to bank data.

An Italian bank used synthetic data to validate a credit-scoring machine learning product from a third-party startup, Ebert says.

Synthetic data could also help with internal development projects. Thresholds set for developers in banks to work on data can be prohibitively high. "It's insane how little access they have," Hazy's Keen says. One bank was unable to run an internal hackathon because its technologists could not satisfy the bank's own data governance requirements.

Erste Group built a retail banking app using Mostly AI's synthetic data, road-testing how the app would work with customers both in terms of its design and in load testing the app's capacity.

In a paper on the work done at American Express, the firm's data scientists say publishing synthetic datasets would help industry innovators develop, train and test machine learning models in areas such as fraud detection.

Land of the hard

Making up data is no panacea, of course. Part of the value of information—especially for investors—is that it's timely. Training generator-models to synthesize data and then synthesizing it can interfere with the speed at which new information becomes available.

Meanwhile, customer behaviour can change quickly, Davis at Facteus points out, as was the case during the Covid-19 pandemic. Forecasting or investment models based on old data can quickly become redundant.

And even fake data generators must give up some fidelity. Ebert says Mostly AI uses 99% of the original data for training to ensure the model doesn't pick up on extreme outliers and inadvertently learn to reproduce them like-for-like.

That said, practitioners think the idea will catch on.

In future, buy-siders may well work with data that's not real. And banks, in faking data, could find a way to monetize information they own but might struggle to use.

It's early days. "This is still in the land of hard," Lucini at Accenture says. Verifying that synthetic datasets are a good enough statistical match to the original will continue to be a tricky problem, he says.

But Lucini thinks the approach will one day become mainstream. Give it three years, he says, and anonymization through synthesizing data will be common practice. wt

Photo by Phil Mosley on Unsplash

Swedish bank finds Covid-19 recovery insights in alt datasets

High-frequency data such as human mobility data and plastic shipments can help investment professionals understand the post-pandemic economic reopening.
By Jo Wright

Jonas Thulin broadcasts weekly video presentations to showcase his predictions for investors and the public. In the early days of the Covid-19 pandemic, he predicted the March 2020 trough in the US equities market, and later, the ensuing V-shaped recovery (economists visualize recoveries and recessions as V's, L's, U's or W's—the shapes they make when charted).

"We were fortunate to publicly call out the equity markets' trough on March 24 last year, following the pandemic sell-off; we were lucky enough to hit the exact day," says Thulin, head of the asset management business at Swedish private bank Erik Penser Bank. "And we argued back in April that based on the data already coming out then that in the third quarter, the US would be above 30%."

Erik Penser is a very small bank, Thulin says, but that prediction helped it punch above its weight and win second place on Bloomberg's ranking of forecasts for pinpointing the US gross domestic product (GDP) growth number.

For asset managers that, like Thulin, use macroeconomic data to understand the economy, high-frequency and alternative datasets to bolster macroeconomic analysis have become crucial in understanding Covid-19's impact on the markets, and now that vaccines are becoming widespread, and some countries have reopened after lockdowns, they're becoming crucial for understanding that also. For these asset managers, traditional measures like GDP just don't get published often enough. Most countries calculate GDP on a quarterly basis; what if you need to understand what is happening in the equities markets this month, or even this week?

Thulin joined Erik Penser four years ago from Nordea, and immediately threw open its strategy and architecture to all asset classes. Rather than running an internal fund, his team's strategy is to allocate assets to external managers in the EU, US, and Sweden, selecting these investments by crunching data—a lot of data.

Thulin says the bank has some 1,600 models and 28,000 time series that it reviews each week. This is, as he puts it, his team's "core view of the world" on which they model all their decisions. "The strength in [using] data over, for example, allocating to an internal fund company is enormous," he tells *Waters Technology*. "Simply put, letting data, products and strategies compete in an objective methodology is key to the future of asset management from our point of view."

This core, however, is only the short-term, high-frequency data, he adds. "When I do the [fundamental] analysis like the backdrop, asking where the business is going, then we are looking at tens of thousands more factors," Thulin says.

Erik Penser Bank has partnered with Sweden-based macro data vendor Macrobond since the latter was founded in 2008, and draws all of its macro data from the vendor. During the pandemic, Macrobond added alternative datasets, such as plastic goods shipments, that Thulin's team have used to make predictions.

Thulin says the thinking goes like this: Analysts can't wait for GDP figures, so they turn to indexes of industrial production to understand economic growth. These, however, are still insufficiently high-frequency, so the next step in understanding how much a country is producing in orders—economists use the US Census Bureau's durable goods orders, for example, to measure industrial activity.

"But then all of a sudden, orders aren't giving you the actual amount of order intake, so you go into shipments. And if you go into shipments, you have to find out what goods are ahead of the curve," Thulin says.

Plastic goods are a good indicator, as orders for these are made early in the production cycles of goods, such as cars, that incorporate them.

"And if you want to understand the core of plastics, you must do the data analysis to see who has the fastest plastics data, and that turns out to be air shipments out of Frankfurt. So that is a great dataset to look at," Thulin says.

Macrobond takes in data from some 2,000 sources, ranging from the traditional (World Bank statistics) to more alternative data from about 350 sources. Many of the newer datasets it has added in post-lockdown days—US airport passenger security checks, the movement of people in Japan, London pedestrian traffic—are data on human mobility.

Macrobond chief commercial officer Howard Rees says an analyst would have to wait a long time to see the impact of the relaxation of Covid-19 restrictions on GDP. "But you can start to look at some high-frequency datasets: how many restaurant bookings are there on [online restaurant reservation service] OpenTable, how many people are visiting different types of shops, and how many people are passing through airports? "We're seeing more and more commingling of the traditional backward-looking, survey-type data and those higher frequency, snapshot-type datasets that tell you what's happening right now," Rees says.

Using Macrobond mobility data, Erik Penser's asset managers calculated the relative strength of the reopening economies of the US and Europe, to arrive at an indicator that explained what was happening in the equities market, Thulin says.

The bank almost exclusively uses Macrobond data, but for very technical instruments, it combines it with Bloomberg data. "If you want to take positions on the US rate curve, for example, there is some pricing data that is quite nitty-gritty," Thulin says. "Or when I have an inflation swap and I want to decide exactly where we want that inflation swap to start and stop, then we use Bloomberg."

Jonas Thulin
Erik Penser Bank

Howard Rees
Macrobond

The firm also runs a sustainability fund for which it uses MSCI data. "We use about 20 million data points to study the sustainability effects of a company, and each company has 100 pages of numerical analysis behind it. When we enter that realm of super detailed, raw quant stuff, we are using complements to Macrobond," Thulin says.

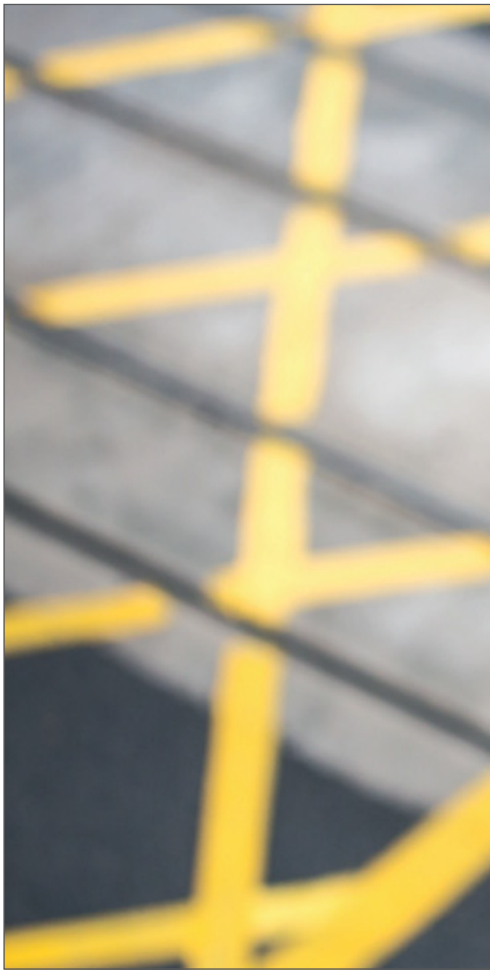
Data differentiator

Any analyst can build models and do math, Thulin says: It's finding unusual data that is the differentiator of asset managers in a competitive market. Clients of providers like Macrobond are thirsty for data, and they are increasingly sophisticated, able to consume huge amounts of it systematically into models and algorithms. While a quant might use a GUI for a sanity check or some quick insight, their main business will be finding alpha in any given content set.

At the same time, clients' migration to cloud providers like Amazon Web Services, and increasingly, Snowflake, has pushed Macrobond to become agnostic in how it sends data to customers, says Greg Haftman, head of data sales at Macrobond. Some clients might want to load the data on-premise, some in a cloud-hosted environment, but however they want it, it has to arrive in the same way and be accessible via a language like Python, he says.

Haftman joined Macrobond in November 2020 from FactSet. In his previous job, Haftman says, he saw how clients were beginning to want to integrate new types of data (audio for example, so they could analyze earnings calls) into relational databases, which struggle with data that is not uniform. Increasingly, clients want to diversify, and store unstructured data and semi-structured data. Macrobond doesn't solely rely on web scraping to ingest data from its sources, but also captures data in semi- and unstructured forms.

Haftman says Macrobond's founders made the decision early in the company's history to store its time series data as Blobs—or "binary large objects," data



such as audio or PDF files or images that are stored as a single entity—in a relational database, rather than in tabular form, as is more common. This decision meant that Macrobond could scale the data it stored and sent to clients. In those days, it had about 6 million time series. When Haftman joined, it had 175 million; it now has 245 million, he says.

"Whatever growth we have, when we onboard new time series, when a user requests data from our server, they request entire objects, which are fully indexed, and that is available via the web API," he says.

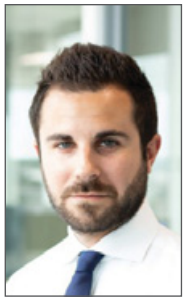
Haftman was brought into Macrobond to help build its new datafeed business, which it announced with the launch of a web API in May. The datafeed is a recognition that consumers need a high volume of data, delivered directly into their operating environments and accessible in statistical applications and programming languages like Python.

"Quants need to have access to data without limitation on what they can consume—in our case, in the number of time series they can consume. How much time does it take to pull out statistics on 2,000 time series, or a million time series? So there are two main ways to deliver data: either you push it or you pull it out; either Macrobond pushes it directly to the client's environment, or the user requests the data on our server. So for us, our API was the strategic decisions we made for a pull mechanism," Haftman says.

He says Macrobond is seeing a growing need for equities and fixed income data in almost real time for multi-asset strategies, and for trying to predict the present, or nowcasting. "So we are trying to get more and more data that has a higher frequency, and also diversify the contributors with alternative data, or data that we haven't really looked at before in terms of investment strategy. That could be data that is not necessarily tied to a particular country, but could be tied to a city or state, for instance," Haftman says.

Discoverability is another difficulty in providing clients with a large volume of data that is constantly updated: How can clients know what's available and then find what they need? If Macrobond had, say, 1,000 time series, that could be listed in a PDF or online, or made available via an interactive portal. But it has 245 million. The company deals with this by categorizing the data in terms of themes, Haftman says. A user could search for a particular sector, for example, and the data is always linked to a country or region.

"The challenge is to make our users and prospects understand that 245 million time series might not mean anything. It's when you start digging into how the data has been categorized into themes and regions, this is how it starts to make sense," Haftman says.



Greg Haftman
Macrobond

Thulin says for the mobility use case, Erik Penser's analysts could search for the term "mobility," or read an article for clues on other kinds of terms that might yield useful data.

"Macrobond also highlights when they add data to their catalogs. The mobility data was one of those points. Our job is to make the math work and say, OK, is this actually explaining something in the market? Is this what we are trading on until the paradigm shifts and we move on to something else?" Thulin says.

Erik Penser Bank has been a client of Macrobond since 2008, and was a client of its predecessor company Ecowin, which was sold to Thomson Reuters in 2005.

While large data vendors like Bloomberg provide macro data, there are some other specialist vendors. Macrobond's main competitors are Refinitiv's Datastream, CEIC, and Haver Analytics. wt

Human Capital

Market structure, e-trading expert Rob Hegarty joins SEC

Rob Hegarty has joined the US Securities and Exchange Commission as a senior policy advisor in its Office of Analytics and Research within the Division of Trading and Markets, to advise on market structure policy and rule-making issues, such as trading, clearing and market data.

Hegarty most recently ran his own consulting and advisory firm, Hegarty Group, and served on the advisory board of merchant bank and M&A advisory firm SenaHill Partners. Before that, he was general manager for financial services at enterprise AI platform vendor DataRobot.

Rob Hegarty

Devi Shanmugham joins Tradeweb as compliance head

Tradeweb Markets has appointed Devi Shanmugham as global head of compliance.

Shanmugham joins Tradeweb from Bloomberg LP, where she was chief compliance officer for the firm's swap execution facility. Prior to joining Bloomberg in 2015, she was assistant general counsel at GFI Group.

Shanmugham will report to Scott Zucker, chief risk and administrative officer at Tradeweb.

Devi Shanmugham

Souvik Das

Taro Kuryuzawa

Souvik Das sails to Clearwater

Boise, Idaho-based investment accounting, reporting and analytics provider Clearwater Analytics has hired Souvik Das as chief technology officer to focus on building out the vendor's SaaS platform.

Das was previously CTO of HR, payroll and benefits app Zenefits, prior to which he was SVP of engineering at healthcare solutions provider Grand Rounds.

Based in San Jose, Calif., he reports to Jody Kochansky, president of product and technology at Clearwater.

KopenTech hires Jill Scalisi to grow CLO platform

Los Angeles-based structured products trading and analytics platform KopenTech has hired Jill Scalisi as chief engagement officer, responsible for business development, launching new products, and overseeing the platform's growth.

Scalisi was most recently co-head of the structured solutions group at Academy Securities, prior to which she spent a decade running her own skincare products company.

Based in New York, Scalisi reports to the vendor's board of directors.

State Street names Taro Kuryuzawa Japan country head

State Street Corporation has appointed Taro Kuryuzawa as country head for Japan, effective immediately. Based in Tokyo, Kuryuzawa will report to Mostapha Tahiri, chief executive for the Asia-Pacific region.

Kuryuzawa joined State Street in May from Deloitte Tohmatsu Group, where he was most recently head of institutional investor coverage in Japan and global lead client service partner.

Exegy taps fintech vet Craig Schachter as CRO

St. Louis, MO-based low-latency data appliance and analytics vendor Exegy has appointed Craig Schachter as chief revenue officer, a new role responsible for setting and leading the vendor's go-to-market strategy.

Schachter was most recently a managing director at SS&C Technologies, responsible for relationship management for the former DST business line, prior to which he was global head of Finastra's fintech ecosystem.

Based in New York, he reports to Exegy CEO Jim O'Donnell.

Theta adds Paul Flanagan as strategic sales advisor

Theta, a provider of buy-side trading technology, has appointed Paul Flanagan as strategic sales advisor, managing relationships with asset managers, hedge funds, fund managers, and pension funds. Theta is due to launch its buy-side trading system, Apollo, later in 2021.

Flanagan has worked in fixed income sales for over 20 years and was most recently managing director for institutional investor sales at Lloyds Bank in London.

Boosted.ai adds Hossein Moein as head of data infrastructure

Boosted.ai, a distributed machine-learning platform for investment professionals, has hired Hossein Moein in New York as head of data infrastructure.

Moein will be reponsible for enhancing data ingestion, analysis, normalization, and post-model output and explainability analysis, as well as developing and managing solutions that give end-users a deeper understanding of the firm's AI offerings.

SIMCORP PROMOTES KROMANN TO CEO

SimCorp has announced that Christian Kromann will succeed Klaus Holse as chief executive officer. Christian has served as SimCorp chief operating officer and member of the executive management board since August 2019.

Klaus Holse will remain a member of the executive management board for the rest of 2021, then serve as a senior advisor until the end of the second quarter of 2022 to ensure a smooth transition.

Christian Kromann

Prior to joining Boosted.ai, he served as the head of data at Kensho Technologies.

BondLink taps Nasdaq's David Allen for sales

David Allen has joined Boston-based BondLink, a provider of software that helps states and municipalities automate processes relating to bond issuance and disclosures, as vice president of sales.

Allen was most recently head of enterprise sales at Nasdaq Governance Solutions, which he joined in 2017 from AI platform GfK, where he was VP of sales for consumer choices in North America.

At BondLink, he reports to CEO Colin MacNaught.

Transcend appoints Steve Vena as business specialist

Transcend, a provider of analytics, optimization, and automation for collateral, liquidity, and funding, has appointed Steve Vena as business specialist. Vena will be responsible for expanding Transcend's business and product strategy to solve complex operational challenges.

Most recently, Vena served on the business transformation global markets team at HSBC and as Brexit transformation lead at Jefferies.

Symphony's Soulier joins LeadingPoint advisory board

Leading Point has added Severine Raymond Soulier to its advisory board. Soulier will help Leading Point expand its product portfolio and international reach.

Soulier was recently appointed head of Emea at Symphony. She was previously head of Emea investing and advisory at Refinitiv and head of continental Europe investment and advisory at Thomson Reuters.

Craig Schachter

Jill Scalisi

SteelEye names Philip Lemmon commercial director

SteelEye, a compliance technology and data analytics firm, has appointed Philip Lemmon as commercial director, working with CEO Matt Smith and the management team.

Prior to joining SteelEye, Lemmon worked for Abel Noser Solutions, specializing in TCA and other compliance solutions.

Meradia adds Brian Lollar as managing director

Meradia, a consultancy for investment operations and technology, has hired Brian Lollar as managing director.

Lollar will support the firm's full scope of investment operations transformation projects and advance Meradia's entry into multiple industry segments. He joins from Accenture, where he was most recently global SimCorp lead for Accenture Capital Markets.

Kingland appoints Matt Kuo chief product officer

Kingland, a provider of data platform solutions, has hired Matt Kuo as chief product officer.

Kuo joins Kingland from AutoQuotes, where he was also chief product officer. He was previously director of product development and enabling tech at CNN, and strategic programs manager for support communities at Apple.

Ex-TT CIO Mike Mayhew soars to Lightspeed

Equities, options and futures broker Lightspeed Financial Services has hired Mike Mayhew as chief technology officer.

Mayhew most recently ran his own consulting firm, Amyna Technology, prior to which he served briefly as CIO and CTO of USA Sports Gaming. Before that, he spent 12 years at Chicago-based trading software vendor Trading Technologies in various roles.

Databricks taps ex-Bloomberg sales exec Roman Ostrovski

Data analytics provider Databricks has hired Roman Ostrovski as enterprise account executive for financial services.

Ostrovski has spent almost 11 years at Bloomberg in various roles, most recently in enterprise sales for trading systems, selling pricing, execution, risk and OMSs to banks.

In his new role, he reports to Alexandra Mysak, director of financial services at Databricks. wt

Jo Wright

SEC's market data plans sow confusion

Jo ponders some of the important pieces of the regulator's national markets system modernization that remain obscure.

Lately, I've been thinking about buying an automatic watch. The appeal for me is that, often, designers leave little windows in the dial through which you can see the internal mechanisms. Looking at the Securities and Exchange Commission's initiative to modernize how public market data is sent out to consumers reminds me a bit of peering through those little windows: you can see the moving parts, but not necessarily how the whole assemblage works together.

Similarly, there are lots of moving parts in the SEC's rethinking of the national markets system (NMS). The first, the so-called "infrastructure rule", became effective on June 8, kicking off the implementation of a system where vendors called competing consolidators will take raw data from exchanges and disseminate it to consumers. These competing consolidators are supposed to replace the two Securities Information Processors (Sips).

Another moving part is the CT Plan, which the SEC ordered the self-regulatory organizations (SROs)—mainly the large exchanges—to draft, and approved on August 6. The exclusive Sips are currently run by two operating committees and three plans; the CT Plan consolidates those into one plan, which will be run by a single—yet to be created—operating committee. The CT Plan will not, as the former plans did, operate the exclusive Sips, but rather the market data system generally.

The SROs have argued that these two pieces—the infrastructure rule and the CT Plan—were so intertwined and co-dependent that the SEC should consider them as one piece of rulemaking. But the SEC insisted that while they are related and tackle the same concerns, they are nonetheless independent: it never mattered to the timelines of one how the other was progressing.

However, even observers who support the SEC's initiative say they're still confused about how these two aspects work together.

"It's still unclear when and how the market data infrastructure rule will intersect with the CT Plan. There still needs to be a plan that covers the exclusive Sips, even under the infrastructure rule, given that the exclusive Sips will operate in parallel with the competing consolidators," says Manisha Kimmel, chief policy officer at market data infrastructure provider MayStreet, who previously spent two years at the SEC.

So what happens to the Sips now, and how will this affect market data subscribers? We know that once there are competing consolidators, they will run in parallel with the exclusive Sips for 180 days. The Sips will still be governed by their current plans and distribute the same data as they do now, to prevent confusion and keep their costs down. Some of their current duties, like timestamping, will fall to the listing exchanges.

After 180 days, the CT Plan operating committee will decide whether the exclusive Sips should be shut down, and can make a recommendation to the SEC either way. The SEC, in turn, can decide if it wants to accept the recommendation, whatever it might be.

Most of the industry seems to assume that the Sips will be decommissioned. But what few seem to realize is that this is not a given: there's no rule obliging either the operating committee to recommend shutting down the Sips, or the SEC to make good on such a recommendation. In which case, what will that look like, and who will run the exclusive Sips?

Revenue remodel

Another point is that, given that the exclusive Sips will now have fewer obligations, they will cost less to run.

Which leads us to another major uncertainty about how the system will work: revenue allocations. Currently, the plans collect revenue earned from the Sips and distribute it to members according to a formula developed just after Regulation NMS was passed in 2005. The infrastructure rule doesn't require a new formula to be created, but it says the operating committee could create one that better reflects the changing SRO obligations if it saw fit.

There's a lot of interest in the allocations among observers of the SEC's efforts. As Kimmel says, "Given the amount of work that went into creating the last formula, it will be interesting to see if the operating committee decides to tackle this issue."

No doubt, in the coming months some of these issues will become a little clearer. Meanwhile, market data consumers will be hoping that all these moving parts do interoperate elegantly under the surface of the dial, so to speak. wt

SMARTSTREAM

Let's talk about the next wave in AI, Machine Learning & Managed Services

SmartStream's fully integrated suite of solutions and platform services for middle- and back-office operations are more relevant than ever – proven to deliver uninterrupted services to critical processes in the most testing conditions. Their use has allowed our customers to gain greater control, reduce costs, mitigate risk and accurately comply with regulation.

With AI and machine learning growing in maturity, these technologies are now being embedded in all of our solutions and can be consumed faster than ever either as managed services or in the cloud.

Simply book a meeting to find out why over 70 of the world's top 100 banks continue to rely on SmartStream.

info@smartstream-stp.com
smartstream-stp.com