



## Examining the LSE's **SEDOL** Fee Gambit

Under the new fee policy, some of the largest users of the London Stock Exchange's identifier codes could see their Sedol spend more than double, though the LSE says the 'vast majority' of clients will see no increase. Industry groups, though, are not buying the exchange's claim.

## Customise your email alerts with Follow Topics

Our most popular Follow Topics include:

- Data management
- Artificial intelligence
- Management & Strategy
- Alternative data



Get started: [www.waterstechnology.com/following](http://www.waterstechnology.com/following)

**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 Falman 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

**To receive WatersTechnology magazine every month you must subscribe to a WatersTechnology Subscription or a Waters Premium Subscription. For more information and subscription details, visit [waterstechnology.com/subscribe](http://waterstechnology.com/subscribe)**

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.

Published by Infopro Digital Risk Limited. Copyright Infopro Digital Risk Limited (IP), 2020. All rights reserved. No part of this publication may be reproduced, stored in or introduced into any retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the copyright owners.



## Market data and consolidation—a never-ending timeline

**No, it's not your** imagination: The list of market data providers is shrinking. And it's a trend that will continue. According to Bob Iati, managing director of TP Icap-owned research firm Burton-Taylor International Consulting, Bloomberg and Refinitiv controlled more than half of the market data pie in 2020 with 34% and 19% shares, respectively. The merger of S&P Global and IHS Markit will claim another sizeable slice, with Moody's, Factset, Morningstar, Ice Data Services, and Six Financial Information accounting for most of the rest.

After that, an enormous patchwork of smaller vendors of all shapes, sizes, and specialties are scrambling to grab fractions of a percentage point of the overall market.

This brings us to Exegy and Vela Trading Systems. As Max Bowie wrote in May, the two have closed a deal to merge under the Exegy brand. Exegy CEO Jim O'Donnell explained that the structure of the deal will help "set us on the path of potential further acquisitions that expand our market data business."

And Iati said that he believes more M&A will happen in the near future and that he wouldn't be surprised if the Big Tech providers—namely Amazon Web Services, Google, and Microsoft—were involved in future deals.

To boil it down, you can either pull your services together to create something robust and long-lasting, or go it alone, and risk going out of business or being poached by a predator, resulting in less-than-favorable outcomes for current employees.

As I've written previously, the key when it comes to M&A and tech development is creating a sticky ecosystem—from front-office trading to back-office operations—and supplying that one-stop data shop. But how might this look in action?

In recent press reports, FactSet was rumored to be both a potential acquisition target and an acquirer, with a FactSet-MSCI tie-up cited by some. Perhaps a deal with Morningstar—and this is pure speculation on my part—would make sense, as the combination would provide solid buy-side, investment banking, and investment advisor coverage, and the FactSet folks are certainly intelligent enough to be able to utilize Morningstar's assets elsewhere across their business. Or, maybe it's a smaller deal for someone like Wall Street Horizon, whose events data could potentially work well with FactSet's Revere Business Industry Classification System, which is its suite of taxonomy products.

While these deals are strictly hypothetical—I have no evidence that any of them are imminent—they serve as illustrative examples of how the M&A cycle could continue to heat up as we head into a post-Covid world.

As always, I'd love to hear your thoughts: [anthony.malakian@infopro-digital.com](mailto:anthony.malakian@infopro-digital.com). **wt**

**Anthony Malakian**  
Editor-in-Chief

# waterstechnology



# Contents

1 **Editor's Letter**

4 **New Perspectives**

14 **Financial institutions battle cyber threat info overload**

Cyber threat intelligence is crucial for the defense of an organization's network, but financial firms have to figure out how to make sense of all the data first. By Jo Wright

18 **Credit Suisse and the Wild West of synthetic prime brokerage**

Industry insiders describe a frontier business with few rules—and plenty of questionable practices. By Kris Devasabai, Helen Bartholomew and Costas Mourselas

24 **Despite client concerns, LSE's new Sedol fees take hold**

Under the new fee policy, some of the largest users of the LSE's identifier codes could see their Sedol spend more than double, though the exchange says the 'vast majority' of clients will see no increase. By Max Bowie



**24**

**COVER STORY:**

Despite client concerns, LSE's new Sedol fees take hold  
By Max Bowie



**14**

Financial institutions battle cyber threat info overload  
By Jo Wright

**18**

Credit Suisse and the Wild West of synthetic prime brokerage  
By Kris Devasabai, Helen Bartholomew and Costas Mourselas



## 28 **Swipe left: repo reporting no match for Brexit, or collateral**

While it enjoyed a happier start than Emir, the SFTR's honeymoon is over now that trades report separately in the UK and the EU. By Samuel Wilkes

## 32 **In fake data, quants see a fix for backtesting**

Traditionally quants have learned to pick data apart. Soon they might spend more time making it up. By Rob Mannix

## 38 **Human Capital**

## 40 **Jo Wright: The ESG Holy Grail doesn't exist ... yet**



### 6

Vendors add surveillance support for WhatsApp alternatives  
By Hamad Ali



### 40

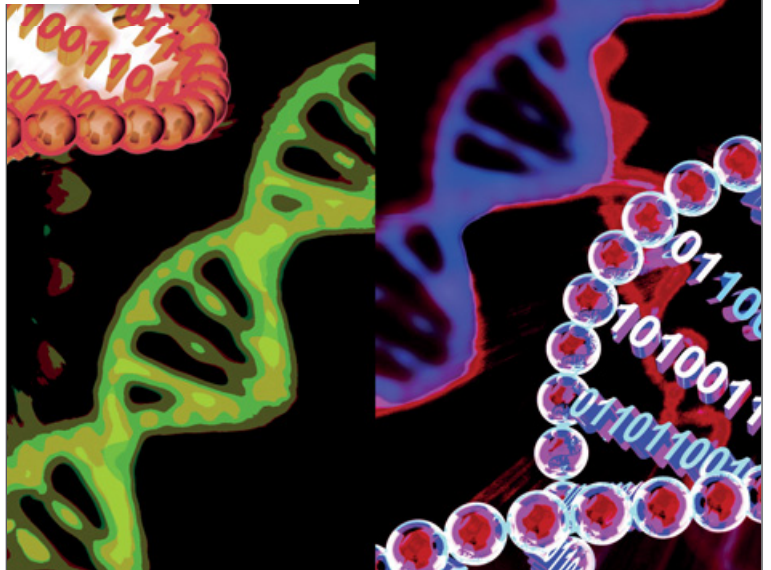
Jo Wright:  
The ESG Holy Grail doesn't exist... yet

## GET MORE ONLINE

News. Webcasts. Video. Audio. Special Reports.  
Get it all at [waterstechnology.com](https://www.waterstechnology.com)

### 32

In fake data, quants see a fix for backtesting  
By Rob Mannix



### 28

Swipe left: repo reporting no match for Brexit, or collateral  
By Samuel Wilkes



# SEC overstepping authority in NMS plan, Nasdaq claims in court

Nasdaq's counsel argued that the regulator does not have the power to give more votes to non-exchange organizations in the Sips' operating committees. By [Joanna Wright](#)

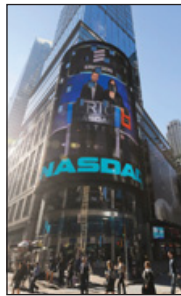
**A** lawyer for Nasdaq told a court of appeals in May that the US markets regulator is overstepping its authority in ordering the biggest exchanges to come up with a new governance plan for the public consolidated feeds of US equities market data, which would give more voting power to broker-dealers and other organizations that sit on the operating committees governing these feeds.

In February, Nasdaq, the New York Stock Exchange (NYSE) and Cboe Global Markets filed petitions in the Court of Appeals for the District of Columbia Circuit, seeking to vacate a plan by the Securities Exchange Commission (SEC) that would change the governance of the consolidated tapes currently operated under the National Market System (NMS)—the Consolidated Tape Association plan, administered by NYSE, and the Unlisted Trading Privileges (UTP) plan, operated by Nasdaq.

Thomas Hungar, a partner at Los Angeles-based law firm Gibson Dunn & Crutcher, said the SEC has no power to give voting rights to organizations that are not exchanges. He said that, via the Securities Exchange Act, Congress gave the SEC the authority to directly authorize only exchanges (referred to as self-regulatory organizations, or SROs) to act jointly in operating the NMS.

“The commission is essentially arrogating to itself the power to transfer authority to Congress via the SROs to non-SROs under the guide of a provision [in the act] that refers only to SROs. It makes no sense, and it violates the structure and purpose of the act,” Hungar said.

The SEC said in 2020 that Rule



Nasdaq says new rules are an appropriation of proprietary data

608 of Regulation NMS authorizes two or more SROs, acting jointly, to file with the commission an NMS plan or proposed amendment to an effective NMS plan. And that is exactly what the SEC then ordered the SROs to do.

This order was part of wider effort by the commission to—as the SEC describes it—modernize how market data is disseminated to US consumers. It is concerned that the two feeds of top-of-book quotes, consolidated from trading venues and published by the Securities Information Processors (Sips), are no longer suitable for the needs of market participants in increasingly high-tech markets. So in December 2020, it finalized a rule that expands the definition of the data to be distributed by the Sips, and opens the door for the emergence of multiple, competing Sips.

At the same time, the SEC looked to modernize the three NMS plans that govern the two Sips. This was the context for the order being disputed in the appeals court: The commission told the SROs, which sit on the Sips' operating committees, to file a new NMS governance plan. The SEC believes the SROs have a conflict of interest, and claims that the Sips feeds are often slower and contain less information than the proprietary market data feeds offered by the exchanges, which control much of the voting power for the NMS plans.

The SEC also says the SROs have a disproportionate amount of voting power on the Sips' operating committees, and the new plan would take some of that power away and give it to other bodies that sit on these committees.

Nasdaq said in a comment letter filed with the SEC last July that the radical revisions to the NMS are unsub-

stantiated, and that the new rules are tantamount to government appropriation of exchange proprietary data. In court, Hungar argued that Congress gave the SROs more authority over the NMS plan as a kind of quid pro quo, since they are subject to unique obligations, like investor protection and direct oversight by the SEC.

“None of this applies to the individual representatives that the SEC is trying to shoehorn into this body,” he said of the Sips' operating committees. “The gerrymandered vote dilution scheme that the commission has imposed on the SROs also violates the statute,” he said.

The SEC's rationale is also discriminatory, Hungar said, in that it wants independent administrators to run the Sips, but it would allow data vendors to serve as those administrators, “even though they face precisely the same alleged conflict of interest that the commission attributes to the SROs,” he said.

Tracey Hardin, assistant general counsel at the SEC, told the court there is nothing in the Exchange Act that indicates that Congress intended to rest sole decision-making authority of the financial markets with the SROs. She argued that Nasdaq looked at the act “in a vacuum,” and said that the language used by Congress in the Act outside of the specific sections cited by Hungar showed that non-SROs—issuers, investors, broker-dealers, and Sips,—were always intended to be involved.

Hardin also noted that Nasdaq did not deny the obvious conflicts of interest that occur when the administrator of the Sips is operating both the public datafeeds and selling its own proprietary data products. [WT](#)



# OpenFin outgrows container tech origins with **Workspace UI** launch

New interface will standardize notifications, user interactions and content presentation. By [Rebecca Natale](#)

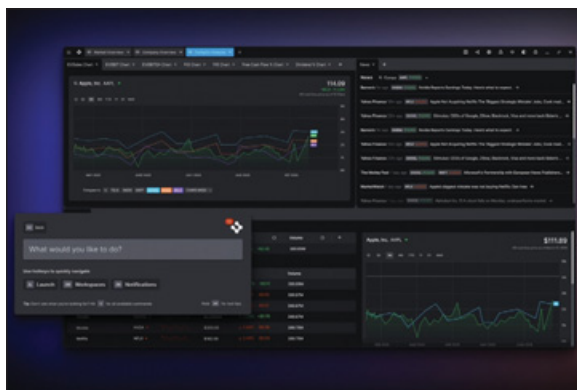
**O**penFin, a New York-based provider of desktop technology, is rolling out Workspace, a user interface comprising a browser, integrated notification center, app store, and digital assistant. The new release will function as a homepage for financial desktops, allowing trading and operations professionals at banks and buy-side firms to launch apps on the desktop.

CEO Mazy Dar says this release is the first time OpenFin has built an out-of-the-box user interface. Typically until now, its customers have built their own UIs on top of the OpenFin container, which can take up to two years. But there has been growing demand from banks and asset managers for the vendor to provide an OpenFin UI, he says.

“What has happened over the last few years is that several of the largest ones who have been building this have said, ‘Look, we have built something, but it’s harder than we thought it was going to be, and it’s taking longer than we thought it was going to take.’ And many of the other firms just have not had the resources to even get it going,” Dar says.

Workspace has been in private beta with five clients over the early months of 2021. Ahead of its April launch, it was already in use at two major banks, two large hedge funds, and a major sovereign wealth fund.

OpenFin began this latest iteration of Workspace about a year ago, as the Covid-19 pandemic was unravelling across the globe, forcing workforces—including traders and ops teams—to work remotely with fewer screens at their disposal. The vendor spent that year collaborating with existing cus-



OpenFin's new Workspace dashboard

tomers to build an interface that could maximize productivity, enable faster decision-making, and leverage existing interfaces into which clients had put significant time and money.

The process led to one of the key components of Workspace, dubbed “Home”—a keyboard-driven digital assistant and app launcher that was originally built as a point-and-click solution before being redesigned. Controlled by keyboard demands, it is prompted by pressing Ctrl+Space, and can search and discover available applications and navigate proper communication channels through its integration with major chat platforms such as Slack, Symphony, and Microsoft Teams. For instance, if using it to peruse contacts and send a message, Home is designed to sense the primary chat platform used by any individual and lists its symbol near their name.

A browser feature focuses on eliminating data silos and distractions. Content from multiple internal and third-party sources are displayed in common windows, which can house several tabs or be dragged around

screens, similar to a traditional internet browser. Workspace’s standardized notification center uses a color-linking scheme to differentiate between internal messages and messages from dealers, vendors, news sources, customers, and others, with the aim of helping users better consume and manage notifications from disparate sources.

“As an application provider, your goal is to be embedded in your customer’s environment so that they’re most likely to see your notifications when they come in,” Dar says. “And that is a problem that no bank, no buy-side firm, no vendor can solve on their own. For that, you need this standardized layer that everybody can build on top of that, which allows all of this to be normalized.”

OpenFin is no stranger to developing standards solutions. In 2018, the vendor led the way on an industry-wide initiative known as FDC3, which aimed to create open interoperability standards for the financial community. More than 20 capital markets firms participated in its inception, and the protocol has since been contributed to the open-source government framework of the Fintech Open Source Foundation.

However, “this is much, much bigger than FDC3,” Dar says, in reference to the standard UI. While interoperability has become an important tool in unifying the user experience, he says standards among interactions and organizing content within the tech—such as launching applications, conducting searches, sifting through notifications, or presenting content in a uniform fashion—do not fall under the scope of the common interoperability language, FDC3. [wt](#)



**Mazy Dar**  
OpenFin

# Vendors add surveillance support for WhatsApp alternatives

Concerns about data leakage have driven some users to rival privacy-focused messaging apps like Telegram and Signal, as WhatsApp policy changes came into force in mid-May. By [Hamad Ali](#)

A controversial update to WhatsApp's privacy policy is driving interest from trading institutions in rival messaging platforms such as Signal and Telegram, and surveillance providers say they are adding integrations to their offerings in response.

"People are concerned about WhatsApp because of the Facebook data-sharing requirement," says Lee Garf, general manager for financial markets compliance at Nice Actimize.

In January, WhatsApp effectively told its more than two billion customers that they had to agree to share data with parent company Facebook by February 8 or WhatsApp would delete their accounts. Millions of users left the platform in protest, migrating primarily to Telegram and Signal instead. Telegram told the *New York Times* that it added 25 million users in just three days in January, pushing it to more than half a billion users, while Signal added nearly 1.3 million in just one day, and was the top free app on the iOS App Store and Google Play Store.

Amid the furor, WhatsApp pushed back the deadline to May 15. The company insists that the update changes little for users, that the controversy surrounding the update was based on misinformation, and that it has shared metadata with Facebook since 2016. However users are still concerned that the tech giant, which has a dismal record on privacy issues, will mine and monetize their data.

Garf says that customers used to be interested primarily in surveillance capabilities for WhatsApp and WeChat. However, since the WhatsApp



“If WhatsApp shares data with Facebook, you have to ask to whom Facebook might sell that data.”

**Neil Bond, formerly of Ardevora Asset Management**

announcement, he has seen more demand from buy-side and sell-side firms for Signal and Telegram.

Neil Bond, former head of trading at Ardevora Asset Management, says firms must prevent information leakage to competitors, and to Big Tech companies like Facebook or Google.

"If WhatsApp shares data with Facebook, you have to ask to whom Facebook might sell that data," he says.

Telegram does not offer end-to-end encryption by default. Signal, which markets itself as a privacy-focused app, does do so, as does WhatsApp, which

in fact implemented the same protocol that underpins Signal. However, unlike Signal, WhatsApp collects metadata on users, including phone numbers, the make and model of their handsets, IP addresses, and any financial transactions made over the app.

The problem for compliance managers at financial institutions is that, while many of these firms ban or limit the use of chat apps and encourage using single, secure platforms to interact with clients, traders like using them. Controlling the usage of WhatsApp has become particularly difficult since the pandemic shifted even some of the most highly regulated staff to remote working.

Nice Actimize saw an uptick in the use of mobile apps over the last year as investment firms have surrendered to staff members' determination to use them, Garf says. Before the pandemic, perhaps 5% of its customers supported



the use of these apps, but that figure has jumped to 50%, he says.

London-based VoxSmart is also developing support for alternative apps, says its CEO, Oliver Blower. “We have begun adding things like Telegram and Signal, which are increasingly popular in the digital asset space, and in emerging markets,” he says.

Blower says he saw an increase in the use of WhatsApp, even before the pandemic, across the buy side and sell side. “The whole market moved over to WhatsApp in 2016. Our clients, who were mainly interdealer brokers had no choice. They had to find a solution, they could not just ban it. So we developed a solution,” he says.

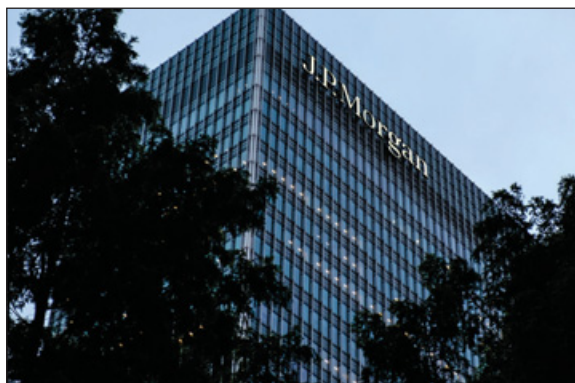
The VoxSmart surveillance solution currently captures three to five client messages per second on WhatsApp.

Garf adds, however, that “firms have taken the position that they will support these mobile applications, but only when they can securely capture them.” There are still some restrictions on how these apps can be used and they need to be fully monitored for breaches in protocol.

JP Morgan banned corporate and investment bank staff from participating in instant message groups involving multiple banks or dealers back in 2013. The bank disciplined senior staff for using WhatsApp at work in 2017, including firing a veteran trader. Deutsche Bank banned WhatsApp on company phones in 2017. However, these firms are starting to bow to the inevitable. In 2020, Deutsche enabled Symphony’s Connect Solution for chat with clients via WhatsApp.

### Surveillance support

Nice Actimize will apply the same technologies and methods it currently uses for surveilling WhatsApp and WeChat to the new apps Signal and Telegram. To capture data from these mobile messaging apps, the user must consent to the vendor installing a monitoring app on their device. Once deployed, the technology operates behind the scenes,



JP Morgan disciplined staff for using WhatsApp at work

monitoring the user’s communications.

The data cannot be captured while it is being exchanged between users since it is encrypted, but once the user grants permission for the app to access data, it can be captured, archived, and analyzed for suspicious activity. “The communications are encrypted, but when you install this plug-in or app on one of the phones, it can get to that information,” Garf says.

Text messages can be analyzed using natural language processing to identify potential misbehavior. Audio files can be converted to text for the same analysis. If the system identifies an activity as possibly suspicious, it will

“The whole market moved over to WhatsApp in 2016. Our clients, who were mainly interdealer brokers had no choice. They had to find a solution, they could not just ban it. So we developed a solution.” **Oliver Blower, VoxSmart**

create an alert for the compliance analyst at a client firm to inspect for possible breaches.

Garf says the surveillance technology operates in much the same manner across all the mobile messaging apps.

“Once it’s captured, and it’s stored in our archive, we then surveil the data that is there,” he says. “At that point, once it is in the archive, the surveillance

is exactly the same” as other applications the company is already surveilling.

Nice Actimize’s surveillance capabilities for Signal and Telegram are scheduled to be rolled out in the second half of this year.

VoxSmart is about to go into production with its surveillance offering for Telegram, but it will take two more months until its tech can support Signal.

Blower says Telegram usage is more widespread in emerging markets. “[It is] mainly in the East, Russia, China, Eastern Europe, and in digital markets, like cryptocurrencies.” However, Signal is still “relatively unknown,” he adds. He says the market demand has not yet fully emerged for Signal, although he says that a lot of clients reached out about the app when the news about the WhatsApp privacy update first broke earlier in the year.

Matt Smith, CEO at surveillance specialist SteelEye, says although WhatsApp has taken a reputational hit due to its new policy, users will not abandon it in enough numbers to cause it concern. “Firms that are progressive will allow it, but they can only do that if they can monitor it,” Smith says.

SteelEye offers surveillance for WhatsApp but is not currently developing support for Signal and Telegram. Roughly 20% to 30% of the vendor’s clients are capturing data from the service, Smith says.

As new communication venues become increasingly popular, compliance teams will come under more pressure to keep track of all the apps regulated users are downloading and using for work. The problem, Smith says, is that bad actors can easily switch to new communication channels if they are determined to carry out illegal activity. “If you are somebody who is doing something nefarious, and you know somebody just installed something on your phone to monitor WhatsApp, what will you do if you want to continue doing the things you do? You just jump for the next thing,” Smith says. **wt**

# Ping An Asset Management zooms in on NLP models for sentiment analysis

The asset management arm of Ping An Insurance (Group) Company of China is enhancing its NLP models to solve complex, non-linear challenges such as overfitting. By [Wei-Shen Wong](#)

**P**ing An of China Asset Management (PAAMC) in Hong Kong—the asset management arm and a wholly-owned subsidiary of China’s largest insurer, Ping An Insurance (Group) Company of China—is looking to upgrade its natural language processing (NLP) models, particularly to account for Chinese sentiment analysis.

Chi Kit Chai, head of capital markets and chief investment officer at PAAMC, tells *WatersTechnology* that vectorization of different words is the key to the success of any NLP algorithm. Vectorization is a methodology in NLP that maps words or phrases to a corresponding vector of numbers to find word predictions, similarities, and semantics.

He says that it can be harder to break down Chinese words in a meaningful way, but PAAMC aims to do so with contextual information and sentiment analysis.

Breaking down Chinese characters is something Boston-headquartered PanAgora Asset Management has also dealt with. The asset manager developed its own machine-learning models to track chat and blog conversations in Chinese to determine market sentiment.

In 2019, Mike Chen, director of equity and head of sustainable investments at PanAgora Asset Management, told *WatersTechnology* that its solution relies on an entire corpus used to train the NLP model of languages to track conversations. The challenge is less in the use of different languages and more in the use of slang or other words in markets-related conversations.



“Our framework merges the generation of alphas and alpha-weighting algorithms using machine learning techniques.”

**Chi Kit Chai, Ping An of China Asset Management**

PanAgora deals with slang words in the Chinese internet community by waiting for them to gain prominence before it updates its NLP library.

“The library is the natural language processing model. It just keeps on updating. When a new cyber slang word gains prominence, if the algorithm sees it sufficiently enough times, it will pick up on it. It’s fully automated and self-updating,” he said.

The work Chai and his team at PAAMC are doing to better analyze Chinese sentiment analysis will go into the firm’s overall machine-learning framework. The framework combines deep-learning neural networks, gradient boosting machines, and advanced regression models.

Chai says in artificial intelligence terms, these combinations are called “ensemble methods.” The framework contains non-linear models, which help PAAMC capture factor interactions and non-linear patterns hidden in alpha signals. On top of that, it provides low correlations among multiple models that can further increase Sharpe and information ratios—measurements that help investors determine the risk-adjusted returns of a security or portfolio.

“Our framework merges the gen-

eration of alphas and alpha-weighting algorithms using machine learning techniques. For factor interaction and non-linearity, an example is the leverage of a company. Linear relations can only assume a company's performance is proportional to its debt ratios. As a matter of fact, debt ratios bear non-linear patterns with a company's performance," he says.

It uses historical structured and unstructured data—including news, price movement information, macroeconomic inputs, and company-specific accounting information—to train its AI algorithms. It monitors more than 300 factors and selects between 20 and 50 factors to construct its portfolios every month.

According to Chris Vera, associate director at asset and wealth management consulting firm Shoreline, using non-linear models means there is a greater ability to incorporate multiple variables to draw conclusions. A linear model would incorporate perhaps two or three inputs to get an output. "Something like, based on these two things, the stock will go up, risk will go down, for example," he says.

In contrast, non-linear models can be used to describe text. "When you and I talk—the sentences we send to each other—we need to put long non-linear formulas to describe [the conversation] because we can use lots of different words and we can construct sentences in different ways. It's more complicated than sending each other numbers because words are quite difficult to describe. There's context, there's language, there's tone, there's volume, there's dialect," he says.

### Building on knowledge

The asset manager, which manages over \$440 billion of assets, is able to leverage technologies and, perhaps more importantly, ideas from all the other units that sit under its parent company. Ping An Group has three main business segments—insurance, banking, and investment—all of which



Ping An Group is headquartered in Shenzhen, China

are supported by its technology arm.

While the asset management business benefits from applying technology that has already been developed in other areas of the group, Chai says different problems require varied solutions from multiple application domains.

For example, Omni-Sinitic—the machine-learning framework that the group developed, which has in the past bested companies like Microsoft, Google, Alibaba, Huawei, and Facebook in the General Language Understanding Evaluation (Glue) benchmark that is used to evaluate natural-language understanding systems—is useful for NLP problem-solving.

For PAAMC, the focus is different. "Here we focus on solving problems in finance and investment. We deal with NLP in our machine-learning framework from a different perspective as we face different challenges. We also put a lot of emphasis on dealing with overfitting when we process very noisy data to extract high-confidence alpha signals," he says.

Overfitting occurs when a model learns the detail and noise in the training data, to the point that it affects the model's performance on new data.

As Shoreline's Vera puts it, overfitting is a stumbling block that happens in data science "when you go from walking to running, and then you trip,

scrape your knee, figure out what you did wrong, and then you start to run again."

Overfitting tends to happen when data scientists throw too much compute power at a model, he adds. "This is where you need to take a step back because you can't just train a lot of data, come up with something that is like a complicated jigsaw puzzle piece and assume that jigsaw puzzle piece can be used for other hypotheses. That's overfitting," Vera says.

According to Vera, PAAMC seems to have achieved "machine-learning sophistication" ahead of other asset management firms. "They're well beyond the use-cases of forecasting liquidity, forecasting changes in risk, formulating portfolio construction—that's all linear. When you're dealing with overfitting, you've moved on to non-linear, and non-linear problems are a lot more data-hungry; they're a lot harder to explain. If you've gotten to the point of overfitting non-linear, then you've been on non-linear for a good amount of time," Vera says.

This could be due to how PAAMC leverages its parent company's technology expertise.

PAAMC's NLP models use news data and corresponding sentiment scores to rank different stocks. As for the overfitting challenge, Chai says machine learning has different solutions to handle overfitting, including cross validation and regularization. "We also use cross-market validations," he says.

In terms of alternative datasets, PAAMC uses Chinese texts from different media, which Chai says provide signals that are robust and that have low correlations to other alpha streams it has. These streams include fundamental, macro, and price data.

Chai says the robustness of the NLP signals depends heavily on the robustness and sophistication of the models. "It is something we spend a lot of time on to differentiate ourselves from others," he adds. [WT](#)



# Asset managers look to raw NGO data for ESG insights

Data from non-profits can be combined with ESG ratings for more bespoke investment insight, investment professionals say. By [Mariella Reason](#)

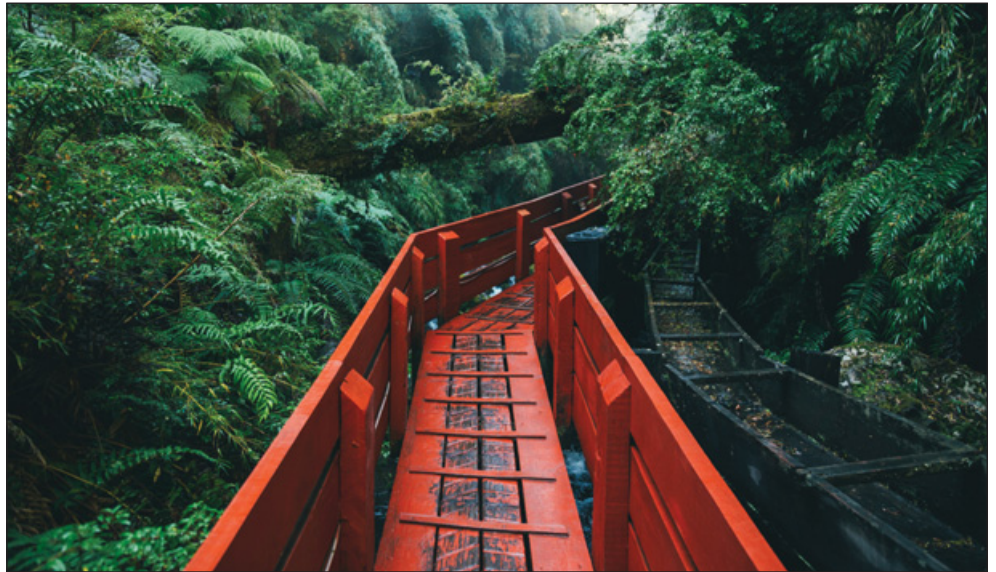
Some investment professionals are turning to the raw data collected by non-governmental organizations to inform their opinions on corporate ESG (environmental, social, and governance) performance, rather than relying on black-box ratings from index providers.

“We are more interested in raw data than we are in ESG ratings. We would rather draw our own conclusions based on raw company data that is either disclosed by the companies themselves or collected by third parties like NGOs,” says Mason Gregory, associate director of investment solutions at MFS Investment Management.

Providers like MSCI and Sustainalytics offer ESG ratings, which are intended to help investors measure a company’s resilience to ESG risks. These providers use proprietary methodologies to identify companies that score well on ESG exposures, and those that do not. These methodologies, of course, are not public, and each provider inevitably encodes certain biases in the ratings, weighting different ESG factors in particular ways according to their own priorities, says Gregory, who specializes in ESG and sustainable investing at MFS.

Relatively raw data from NGOs allows investment professionals like him to make those decisions for themselves, and apply their own weights.

NGO data can be gathered directly from an NGO itself, which can be a time-consuming process, or from a data provider that compiles reports from a number of different NGOs. Ratings providers themselves draw on NGO data. Data from NGOs like



“We would rather draw our own conclusions based on raw company data that is either disclosed by the companies themselves or collected by third parties like NGO.” **Mason Gregory, MFS Investment Management**

Global Fishing Watch is accessible in Bloomberg’s Professional Terminal. Last year, Refinitiv signed a contract with Sigwatch, a UK-based provider of global NGO and ESG issue tracking and reputational impact data. The collaboration gives Refinitiv customers access to insight into NGO campaigns affecting more than 19,000 companies, brands, and projects, and Refinitiv is enhancing its due diligence reports with Sigwatch data.

Gregory says firms can also gather NGO data from watchdog groups that issue reports on supply chain controversies, such as cocoa supply chains that include forced child labor issues. The NGO Mighty Earth provides a free Cocoa Accountability Map which combines cocoa supply chains from nearly every major cocoa and chocolate company operating in the Ivory Coast.

## Problems with ratings

Mike Chen, director of equity and head of sustainable investments at PanAgora Asset Management, says the firm is in the early stages of looking into how NGO data can be used for ESG analysis. He agrees that rating a company on ESG is an individual, personal, and values-based process. “There’s no right or wrong; it’s like asking someone what their favorite color is. I think therein

lies the problem of using commercial ratings—you're accepting somebody else's decision for you," he says.

Companies that might be controversial to an investor might score surprisingly well on ESG factors with a ratings provider, Chen notes. For example, Sustainalytics gives tobacco company Philip Morris a medium ESG risk rating, while MSCI—which scores ESG from a scale (from worst to best) CCC, B, BB, BBB, A, AA, AAA—gives them a BBB rating, also the middle of the road. Chen says a company might market an unhealthy product like cigarettes, and score poorly on the "S" pillar of ESG, yet score highly on G because it has a diverse board with many women in positions of leadership. The company has also pledged to create alternatives to tobacco cigarettes.

"But does a good governance structure offset the fact their product contributes to the deaths of millions of people? According to me, no," Chen says.

Someone with their own values, life experiences, and goals may see the company in a different way from their peers. "That's why using ratings is very problematic because you are outsourcing your own decision about what matters to you, to other people," Chen says.

Chen says that when collecting ESG data, it's important to not only look at surface-level datasets. "When people think about diversity, people think about gender; that's the most obvious thing. And that data is relatively easily available," he says. But there's more dimension and nuance to gender diversity data than the sum of a person's X and Y chromosomes. Chen says investors need to consider socioeconomic background, nationality, ethnicity, age, and education level to get a different dimension to diversity.

Using raw data straight from the NGOs gives PanAgora the opportunity to draw its own conclusions on whether a company should be included

in a particular ESG portfolio. Chen says that these types of data are more useful to him when compiling an ESG-based portfolio than pre-packaged information from an ESG ratings agency.

"If an NGO puts out a report that a certain company has bad employment practices, or a company has been polluting, we read that data and then decide how important or relevant those reports from the NGO are to the way we view ESG," Chen says.

Both PanAgora and MFS still use ESG rating agencies alongside their own methods of data collection. Gregory says investors can still get raw data from aggregators, but that he finds the actual company ESG ratings less useful, "We often don't agree that the issues driving them are the most material for a particular company," he says.

Chen says PanAgora practices multi-factor investing. A multi-factor model is a financial modeling strategy that draws on multiple factors to analyze and explain asset prices, so Chen can draw information from NGO data, newspaper reports, the company's own reports, and sell-side analyst reports to

“If an NGO puts out a report that a certain company has bad employment practices, or a company has been polluting, we read that data and then decide how important or relevant those reports from the NGO are to the way we view ESG.”

**Mike Chen, PanAgora Asset Management**

inform the investment process. "Any decision to buy or sell a given company is not based on one information source alone. Then once you have all these raw ingredients, you've got to synthesize it into how you want to weight each piece of information," he says.

Gregory says another advantage of the data that NGOs collect is that it's often unique. Organizations like the Workforce Disclosure Initiative, for

example, collect salary information on global companies that isn't available anywhere else, he says. The WDI aims to improve data on how companies manage workers across their operations and supply chains. In 2020, 141 global companies took part in its annual survey. "A lot of this data is not information that the companies voluntarily disclose, or if they do disclose it already, it's not typically that detailed," he says.

NGOs also collect data regardless of the cooperation of the corporate entities whose ESG performance asset managers want to understand. That removes a lot of the bias involved in corporations' own reporting, which still forms a substantial portion of the ESG data that the buy side relies on.

In addition, ratings from the big providers tend to be "backward-looking," Gregory says, particularly regarding controversy. In 2015, the US Environmental Protection Agency said Volkswagen had violated environmental laws by selling cars in America that cheated emissions control tests. The resulting scandal was dubbed Dieselgate, and it tanked the auto manufacturer's overall ESG rating, Gregory says.

After Dieselgate, MSCI downgraded Volkswagen to a CCC rating, the lowest rating the provider offers, and the company now has a B rating (only one rating higher than CCC). "You can decide to punish a company for what they have done in the past, but as active investors we have an obligation to focus on the future and try to understand how a company is likely to perform going forward," Gregory says.

This is where the results of an ESG ratings agency can be unhelpful for an asset manager looking to see not what a company has done in the past, but how they are positioned for the future.

"We are focused on understanding how ESG performance on various issues is likely to impact the sustainability of a company over the long term from a financial perspective," Gregory says. [WT](#)

# Esma: Firms struggled with trade data porting after CME wind-down

The EU regulator had to coordinate efforts with local NCAs to clamp down on failures made by counterparties to meet Emir guidelines for porting data. By [Josephine Gallagher](#)

In May 2020, the Chicago Mercantile Exchange announced plans to wind down several of its regulated businesses, including its European trade repository (TR), CME ETR. The rollback meant that thousands of regulated market participants reporting trade data to ETR would have six months to reroute their reporting flows to a new TR ahead of November 30, 2020, when the services were scheduled to be terminated. However, in a report published on April 15, European regulator the European Securities and Markets Authority (Esma) said that some counterparties failed to follow its guidelines on porting data from one TR to another.

“The process that the TRs and the counterparties should follow is described in the guidelines on portability,” said Jakub Brettl, senior supervision officer at Esma, during a press briefing on April 28, in response to questions from *WatersTechnology*. “It’s really important that this process is followed, because that way we can ensure that there is, indeed, no impact on data quality and data integrity.”

Esma’s guidelines for porting were finalized in 2017, and aim to make it easier for reporting entities to move records from one TR to another so that they can easily change providers if they want to, and to establish a consistent way of doing so. Reporting entities must move all their open and historical trade data to their new TR so that regulators in each jurisdiction can access the reports.

The guidelines also say that the new TR should not accept duplicate reports by TR participants, and the old TR



should not accept reports with action types “cancel” or “error” made by TR participants.

However, Brettl said, some counterparties that were customers of CME ETR failed to port their historical data to their newly designated TR. Instead, they sent cancellation messages on their existing derivatives records with CME ETR, intending to start afresh when reporting to their new TR.

If reporting parties do not follow Esma’s porting guidance, Brettl said, it can distort reported data flows, hampering the regulator’s ability to monitor the market.

“This causes, of course, many problems, because if historical data stays in the old TR, it creates certain spurious, fake data flow, if you will, because these cancellations do not represent real economic activity, which is what we want to capture. So, it creates a certain level of noise,” he said.

During the CME porting window, Esma told local regulators (so-called national competent authorities, or NCAs) and the EU TRs to be on the lookout for porting that did not follow its guidelines.

“We have been very clear to the TRs, and we’ve tried to be clear to everyone, that this is not the way to do reporting. Whenever we saw, with the assistance of the TRs, that there might be clients attempting such ways of moving the data, we communicated with the NCAs to get them involved,” Brettl said.

A spokesperson for CME says ETR noticed some isolated cases where clients transferred data using a procedure that did not follow Esma’s guidance. They said that after an investigation, CME stressed the importance to its customers of following the porting procedures during the transfer period.

Esma’s April 15 report examined



the quality of the data submitted to TRs under the European Market Infrastructure Regulation (Emir) and the Securities Financing Transaction Regulation (SFTR). The report says that—in addition to CME’s commercial decision to shut ETR—Brexit also triggered changes in TR market structure that led to a high volume of data being ported in Q4.

Other factors contributing to higher-than-normal activity over the same period included Intercontinental Exchange’s decision last year to wind down its own Trade Vault rather than provide TR services in the EU post-Brexit, and the creation of new TRs to support reporting. For example, the Depository Trust & Clearing Corporation and the London Stock Exchange’s UnaVista established TRs in the EU, and Regis-TR established a UK TR—all of which involved shifting records from their UK to their EU TRs. There are now four TRs offering services in the EU for both Emir and SFTR—the DTCC’s DDRIE, UnaVista, Regis-TR, and KDPW.

### Complex operation

John Kernan, CEO of Regis-TR UK, says all counterparties that moved data from CME ETR to Regis-TR during the transfer window ported all their open and historical trades. However, he says, there were some problems in porting data in the format used by CME to that of Regis-TR or processing old reporting data that predated revisions made in 2017 to reporting practices under Emir. Under these updated regulatory technical standards (RTS), counterparties had to follow set standards for reporting to a new TR, ensuring, for example, that both parties to a trade agreed that the reported data was correct.

TRs use their own proprietary formats for reporting data, Kernan says, which complicates porting. This is expected to change under updates to Emir, known as the “Emir Refit,” which will harmonize reporting stand-

ards by forcing TRs and counterparties to report using machine-readable XML schemas.

The report says that from about 400 million submissions to CME’s ETR in the month of September 2020, its volume sharply reduced. UnaVista and Regis-TR were two of the biggest recipients of reporting flow from the CME ETR unwind, with an increase of more than 150 million submissions between September and December 2020.

In some instances, counterparties were slow to select new providers during the transfer window and had to be prompted by the regulators to act quicker, Brettl said.

“It’s not something that can be typically performed on one day’s notice, and so those cases, those counterparties that we observed that were a bit slow in selecting a new TR we had been actively engaging with the NCAs to then make sure that they, on their side, can engage with those counterparties and make sure that everything goes well,” Brettl said.

Toward the end of the transfer period, Regis-TR’s Kernan says Esma selected several counterparties that had large sums of non-RTS quality data

and allocated them with a new TR. Under Esma guidelines, the regulator can appoint a TR if an old TR has not notified it about a new TR or the new TR refuses to receive the counterparties’ data.

“Esma took that list of clients and allocated them out to the TRs, so there wasn’t a selection process or anything like that on our part. Toward the end of the CME wind-down, each TR received a number of clients with old historical data that wasn’t the requisite standard,” he says.

It took some counterparties nine months to port their data to a new TR, Brettl said during the press briefing. While the May-to-November transfer window given to counterparties was six months, and not nine, Esma was notified of the unwinding prior to the public announcement, says an Esma spokesperson, responding by email to *WatersTechnology*.

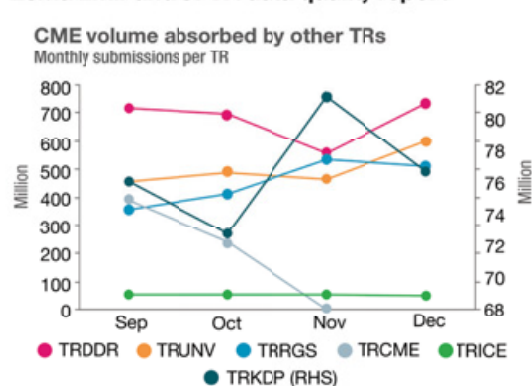
“Esma knew about the wind-down plans well in advance and we had been discussing operational aspects with the TR well ahead of the public announcement. CME’s operations ceased in November, but technically speaking the wind-down was formalized by the withdrawal of the registration by Esma in December,” the spokesperson says.

The official deadline for the wind-down was deemed December 2020, and all open and historical data was successfully ported to other TRs ahead of Brexit, they say.

A spokesperson for CME says that once all clients were successfully offboarded and their “live” data was ported—which was completed by November 30—there were a further few weeks of porting of historical (that is, non-“live”) data from the CME ETR to other TRs. *WatersTechnology* understands that while the client-facing porting process took six-and-a-half months, the overall timeline for the CME TR wind-down from start to finish was closer to eight to nine months.

UnaVista and the DTCC declined to comment for this article. [wt](#)

### Esma Emir and SFTR data quality report



Note: Total number of daily submissions per month and TR. 'TRDDR' is DDRL, 'TRUNV' is Unavista, 'TRRGS' is Regis-TR, 'TRCME' is CME TR, 'TRICE' is ICE TVEL, and 'TRKDP' is KDPW. KDPW is on the right-hand side vertical axis. Figures for KDPW (November 2020) are underestimated with around 6mn due to a technical issue when processing the data files.

Source: Trade repositories and Esma calculations

# Financial institutions battle cyber threat info overload



Cyber threat intelligence is crucial for the defense of an organization's network, but financial firms have to figure out how to make sense of all the data first. By Jo Wright

**T**he Chinese general Sun Tzu said that if you know your enemy and you know yourself, you need not fear the outcome of a hundred battles. If he added that too much information about your enemy can be overwhelming, history does not record that.

Knowing the enemy is no less important to the security operations centers (SOCs) at financial organizations than to Han Dynasty military strategists. But the sheer volume of threat intelligence—the data an organization uses to understand the information security threats it could face, or breaches that have already taken place—is overwhelming cybersecurity professionals. In 2021, one of the main roles of nonprofits and private sector companies that offer threat intelligence is curation and finding signals in the noise for clients and members.

When Anomali, a provider of cybersecurity tools, was founded eight years ago, the volume of indicators—identifiers of threat actors or events, such as IP or email addresses—that firms had to process was manageable, says its CEO, Hugh Njemanze. A customer might have had to deal with a database of about 100,000 risks or indicators, and feed those into the tools in their SOC systems.

“But just a year later, it was a million indicators that we were helping with, and then it was 10 million, and 100 million, and now it’s several billion indicators that are active,” Njemanze says. “It completely overwhelms not just the humans, but any tools

they have. They need to find ways to filter down the volume of threat intelligence.”

Like good generals, chief information security officers (Cisos), too, must understand the risks they face in the internet battlefield, where attackers range from everyday chancers sending phishing emails, to sophisticated spies hired by nation states (known as advanced persistent threats, or APTs). And there is a booming industry of vendors offering platforms and services for threat intelligence: The global industry is projected to grow from \$10.9 billion last year to \$16.1 billion by 2025, with security analytics to register most growth, as users want to find patterns



in attacks across network infrastructure, according to market research.

Vendors like Blueliv, ThreatConnect, and Cyjax offer threat intelligence platforms with differing degrees of modularity; Cobwebs Technologies extracts targeted insights from big data; ZeroFox provides threat intelligence across the public attack surface. Many vendors, as Anomali does, provide access to data feeds from providers like Symantec, CrowdStrike, and FireEye, which an organization can pull into its security and event management software (Siem) to collect logs and event data generated by apps and security devices.

So there is no lack of information out there. The problem in 2021 is that there

is too much. As Njemanze says, Siem systems generate a huge volume of alerts. Not only is this tiring for those people who must monitor them, it's also difficult to detect actual security threats. The problem is compounded by a shortage of qualified staff at many institutions. In a survey of cybersecurity professionals conducted last year, 31% said understaffing was the biggest challenge they faced; 18% said it was keeping up with the volume of security alerts.

### Filtering solutions

Njemanze says vendors use a number of approaches to filter data for clients. "You could filter by certain classes of intelligence, so maybe you only care about IP addresses, or compromised files. We have built algorithms, and we have a dedicated lab of practitioners. They look for what is happening in real time, what is trending now, what has gone dormant," he says.

"So we very aggressively curate our repository of threat intelligence to deactivate threats that are no longer active, to promote threats that are currently active, or targeting a specific vertical," Njemanze says.

Anomali's Match tool matches large volumes of threat intelligence in a client's Siem software with activity from its network. Siem software collects event logs, and Match compares that log to the Anomali database of threat intelligence, and returns the information. For example, it could show that there were six sightings of a particular bad actor in a particular month, the actor was first seen in this system, and the attack spread to another system. "Otherwise, you just have a repository of threat intelligence dumped on you, and it's very hard to find something that can consume it. That's why people tend to filter out the threat intelligence," Njemanze says.

This backward look is important, he adds, as most attacks go unnoticed, sometimes for years.

### Sharing is strength

Financial firms don't just have to rely on the private sector for threat intelligence: they can also join an Information Sharing Center (Isac). Isacs are nonprofits organized by industry verticals that act as forums for the sharing of threat intel-

**“We very aggressively curate our repository of threat intelligence to deactivate threats that are no longer active, to promote threats that are currently active, or targeting a specific vertical.”**  
**Hugh Njemanze, Anomali**

ligence. Finance has its own Isac, called the FS-Isac, which was founded in 1999, and has a membership of thousands of financial firms in 70 countries.

"The memberships of these Isacs feed information up to the organization as they recognize threats or attacks against their systems, and the community then shares it back out. It's like a neighborhood watch concept, and it's pretty powerful," Njemanze says.

The deputy Ciso of a New York-based investment firm that is an FS-Isac member says they believe it's important to subscribe to an Isac.

"We benefit not just from learning about active threats, but also getting advice from other people in the same industry. You can say, 'Hey, we have this problem, how did you deal with it? Is this something you are concerned about?' Having that community of other security professionals who have dealt with similar situations is helpful," they say.

Having this community on a platform like FS-Isac's is important because these provide different levels of anonymity, they add. "Having it around a service gives you that comfort of nondisclosure and making sure there is no issue with sharing. You don't want to share too much, but you also want to share enough that it's helpful to both your own community and others."

One of the FS-Isac's main roles is making threat information meaningful and useful to its members, says Teresa Walsh, global head of intelligence for financial services information sharing at the nonprofit.

The FS-Isac sends out intelligence in a number of ways—via listserv, on special calls and briefings, and through research



reports. It also has a centralized platform for threat intelligence sharing called Share and a recently launched chat app called Connect.

Share uses tagging technology to help users find information organized by category. “Users can share information on different types of cyber threats, how they are impacting their businesses, threats they hear about but don’t know much about, and more general information—how others are responding to Covid,

board could not post market announcements. The exchange shut down trading for the day but was attacked again on reopening. The exchange had to suspend activities every day for four days, halting trading in cash, debt, equities, and derivatives markets.

As Bloomberg reported in February, the NZX had been the victim of a distributed denial of service (DDoS) attack, an unsophisticated but often effective attack that enlists thousands of systems

cant damage after the deadline for their Bitcoin ransoms.

Walsh tells *WatersTechnology* that FS-Isac’s membership gave the organization an advantage—because its members were willing to share their experiences, the body could take in that data, and understand that the extortion emails and DDoS attack were connected; that the unsophisticated hackers were almost certainly not Fancy Bear or Lazarus Group; and it could track the hackers as they systematically targeted firms in first Europe, then the Americas, then Asia-Pacific.

Share generates alerts at various stages of urgency; for instance, if a member reports an incident, or new research is issued, members will see that in their cues. Those alerts accumulate over time, and FS-Isac can analyze them for patterns, as it did while the DDoS extortion campaign unfolded. For ongoing events like that, Walsh’s team can centralize and analyze information and then build visualizations to make it easier for members to consume.

“You do have to tailor it to your audience. People are not going to read massive amounts of data. So you could, for a DDoS attack, include 10,000 IP addresses [of the botnet], but the most important thing is when did the attack happen, what did it attack exactly—your public-facing websites or less visible infrastructure?” Walsh says.

“We started with a simple Excel spreadsheet and just compared the different attributes to the attack we were seeing to understand how the members were experiencing all this, and then sharing it with each other,” Walsh says.

Connect was also useful during the DDoS extortion campaign, Walsh says. “People had a lot of questions about everybody else’s experiences. What type of threat did you get? Were they demanding payment? How did you respond, how did you brief your staff? Did you answer the emails? ... It was almost like a support group,” Walsh says.

Those individual experiences enabled, in the aggregate, a crowd-sourced analysis of events, and the FS-Isac could then reach out to other members to find those that might have experienced similar extortion attempts but had not shared them.

““

**“We benefit not just from learning about active threats, but also getting advice from other people in the same industry. You can say, ‘Hey, we have this problem, how did you deal with it? Is this something you are concerned about?’ Having that community of other security professionals who have dealt with similar situations is helpful.” Deputy Ciso of a New York-based investment firm**

or briefing their board on cyber risks,” Walsh says.

“They use each other as sounding boards for what works and what doesn’t and what is best practice, whether they are a huge multinational conglomerate or a small community institution serving a small customer base,” Walsh says.

One advantage the FS-Isac has is that it comprises thousands of members worldwide that have, over the years, become increasingly comfortable sharing threat intelligence with their peers anonymously. This is not a small thing—for financial organizations, admitting to a security breach can be a massive reputational risk and invite future attacks.

Because of the volume of reporting centralized on Share and, more recently, Connect, the FS-Isac has been able to draw inferences that might have gone unnoticed otherwise and feed those back as insights to its membership.

### Targeted attack

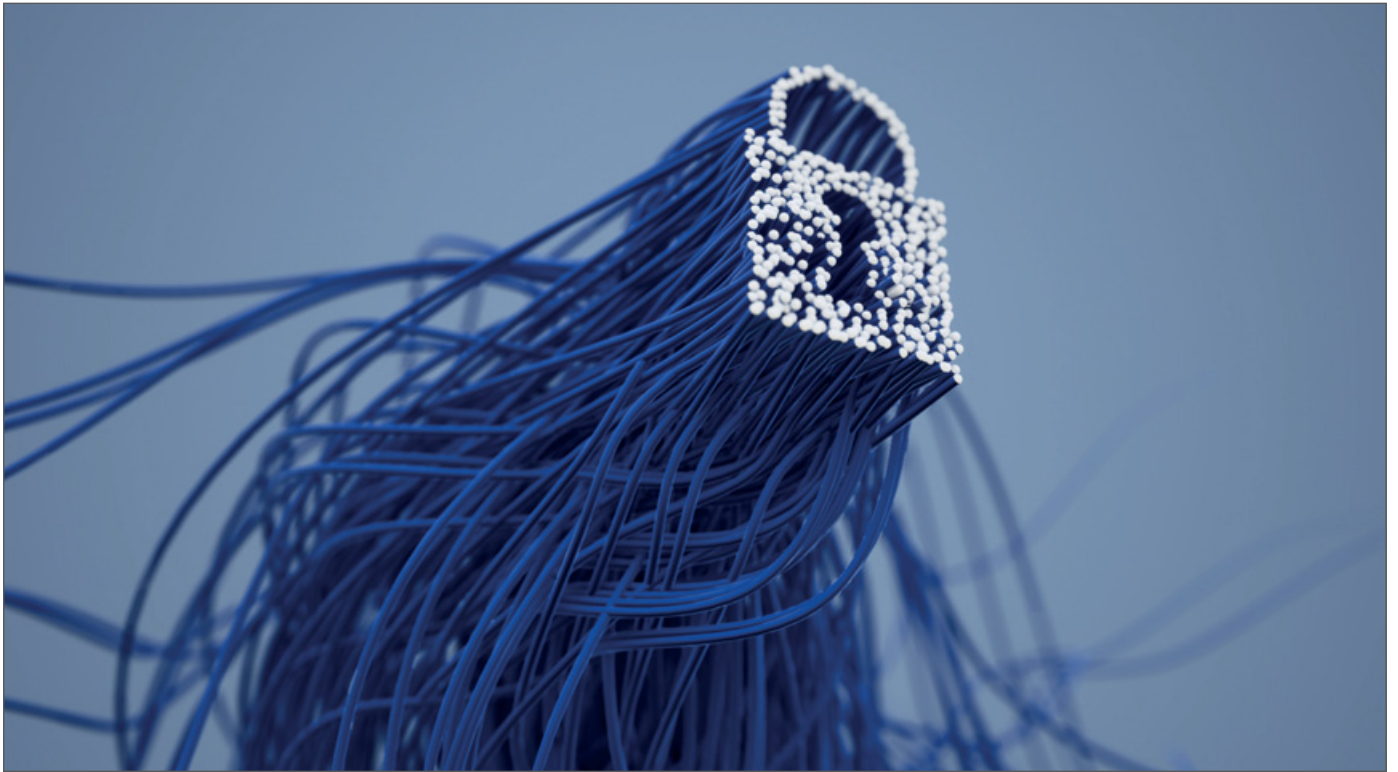
Perhaps the most dramatic demonstration of this capability came in the wake of attacks on the New Zealand Stock Exchange (NZX). In August last year, a massive wave of internet traffic deluged the NZX’s website, slowing it down to the point where the country’s main

infected with malware (sometimes called “zombie armies”) into making requests of a targeted server, flooding it with traffic and shutting it down. The NZX moved its servers out of reach to the cloud, but the malicious hackers moved on to targeting companies listed on the exchange, and to other organizations.

The attackers sent target firms emails in which they claimed to be from notorious APTs like the Lazarus Group and Fancy Bear. The extortion emails threatened more DDoS attacks if they weren’t paid sums in Bitcoin.

At the time these attacks were spreading around the world, it was not clear they were linked. However, as FS-Isac members began to receive the extortion emails, the organization was alerted and began connecting the dots. FS-Isac said in February that its analysis found that the extortion emails and DDoS attacks appeared to follow a similar modus operandi and seemed to indicate that these were connected events that proceeded in a systematic way across the globe.

Apart from NZX, targets ran the gamut from large banks to fintechs, insurance companies, asset managers, and loan companies. The FS-Isac said at the time that none of its members paid the ransom, and none saw signifi-



## Know thyself

To make inroads into all the information at their disposal, firms must look inward as well as turn to outside help. After all, while knowing your enemy is important, as the second part of Sun Tzu's adage goes, you also must know yourself.

Aite Group senior analyst Tari Schreider says many Cisos take a "more is more" approach to threat intelligence. "One tendency I see in financial institutions is where, when flush with cybersecurity investment cash, they go out and subscribe to as many different threat intelligence sources as possible. The intelligence itself is useful, but if you're not interpreting it correctly, you might misread it and end up putting defenses in the wrong place," he says.

A better strategy is to start small and build up slowly to what the organization needs, taking care to understand its special vulnerabilities and what information is actionable along the way.

"To really understand threats, you must take intelligence and enrich that data to compare to your own attack surface and understand where you are vulnerable. Then you risk score the vulnerabilities, and that is the risk to your organization," Schreider says.



**"To really understand threats, you must take intelligence and enrich that data to compare to your own attack surface and understand where you are vulnerable. Then you risk score the vulnerabilities, and that is the risk to your organization." Tari Schreider, Aite Group**

He says it's crucial that the work of analyzing threat intelligence not be confined to the SOC, or siloed in any one department of an organization. The deputy Ciso of the investment firm agrees, saying communication with the business and compliance functions is probably the most important—and often most difficult—part of their job. Schreider says an organization should have a comprehensive threat intelligence platform into which it feeds intel, but that intel should be contextualized to the infrastructure and the business of the rest of the firm.

While firms may see themselves as individual siloes, hackers do not. "Your attack surface is a bowl of tasty threat vectors," Schreider says. "Hackers view your organization as an attack surface on which you just happen to have front-office technology, back-office

technology, network technology, etcetera. They look past classical organizational constructs and see threat vectors, asset vulnerabilities and exploit potential."

Covid-19 and the shift to remote working made this attack surface even more diffuse, he concludes. "Covid-19 made organizations' attack surface look like Play-Doh and spread it all over the place. Your network endpoints are now your employees working in their homes, they become the weak link in the security value chain. And they have caught the eye of the hackers. Whereas you would have a trader sitting on the trading floor, that individual is now at home trading securities. The same security protocols that exist on a trading room floor just don't exist in a home setting. Hackers know this and will always follow the path of least resistance," Schreider says. [WT](#)



**Tari Schreider**  
Aite Group

# Credit Suisse and the Wild West of synthetic prime brokerage



Industry insiders describe a frontier business with few rules—and plenty of questionable practices.

By Kris Devasabai, Helen Bartholomew and Costas Mourselas

**W**hen a veteran industry lawyer describes part of the prime brokerage business as “a bit of a Wild West,” it might bring to mind a quick-draw duel in the dust of a frontier town, antagonists squinting at each other in the noon sun, hands hovering over their pistols.

That image captures one part of the drama that has catapulted prime brokerage into the headlines. When a clutch of banks discovered a defaulting client—Archegos Capital Management—had similar positions with each of them, Morgan Stanley and Goldman Sachs fired the first shots, selling the hedges that all of the banks were holding. Credit Suisse caught the bullets, and a loss of \$4.7 billion, with more to come.

That’s just one part of the story, though, and it’s not what the lawyer means. Instead, he’s referring broadly to the lawlessness of synthetic financing, where banks provide leverage to hedge funds via swaps—an anything-goes, sheriff-free business, where standard risk management practices seem not to apply.

“You’ll find all sorts of things happening in prime risk management, from sophisticated to not,” says the former head of one Wall Street prime brokerage. “Exposure in prime businesses is not appropriately remunerated. As the Credit Suisse

fiasco demonstrates, the risks are huge and the leverage completely outsized.”

Synthetic financing accounted for 54% of prime brokerage income in the first half of 2020, according to consulting firm Finadium.

With no formal limits on leverage, margin requirements for equity swaps can vary from 5% to 50%, sources say—levels that are set by a wide array of models, some bluntly static, others smarter and dynamic. Critics also point to a lack of transparency, which makes it tough to gauge and manage risk. Trades can be housed in multiple legal entities, with some banks relying





on archaic, pieced-together systems to calculate and monitor their exposures.

“Even within a firm, between different types of equity derivatives, you may encounter different systems, controls and capabilities,” says another former prime brokerage head.

Prime brokerage insiders describe a dysfunctional dynamic where the largest and riskiest clients often push for—and receive—preferential treatment, creating a race to the bottom to win lucrative accounts.

“When you’re a big client—which Archegos was—and prime brokers are fighting over your business, you can push

for more lenient terms in the synthetic world,” says one industry lawyer.

According to *Risk.net* analysis, Credit Suisse agreed to softer terms than Archegos’s other brokers—the default happened before the bank could switch from static to dynamic margin models, a step that is understood to have been planned for the first half of this year. The bank’s internal structure, and turnover within the business, may also have contributed to its losses. And it was reportedly slower on the draw, waiting longer before selling the huge books of stocks it was holding as a hedge for its swaps with Archegos.

“

**“Exposure in prime businesses is not appropriately remunerated. The risks are huge and the leverage completely outsized.” Former head of a Wall Street prime brokerage**

But it’s not just a Credit Suisse story—Nomura’s loss of \$2.3 billion was on roughly the same scale, relative to the exposure held—and industry insiders on both the buy side and sell side paint a picture of a frontier business with few rules and plenty of bad practices.



Some say cost-cutting and a high leadership turnover left the division rudderless.

Regulators including the US Federal Reserve and the Securities and Exchange Commission (SEC) are now signaling new rules for prime brokers and their hedge fund customers.

### The good, the bad and the ugly

Prime brokerage can be a risky business at the best of times, let alone when dealing with a client such as Archegos.

The family office—which managed the personal wealth of Bill Hwang, a former hedge fund manager who settled charges of insider trading in 2012 and pleaded guilty to wire fraud that same year—placed big, concentrated bets on a handful of stocks, including media conglomerate ViacomCBS, music streaming

service Tencent Music and online discount retailer Vipshop.

Archegos spread its positions across six main prime brokers—Credit Suisse, Deutsche Bank, Goldman Sachs, Morgan Stanley, Nomura and UBS—and also traded with several others, including Wells Fargo, Mitsubishi UFJ and Mizuho, on a smaller scale.

This is not unusual. After being badly burned when Lehman Brothers defaulted in 2008, hedge funds began dividing their business across multiple prime brokers, leaving each counterparty with only a partial view of their portfolios.

The use of total return swaps further obscured the risks. This type of synthetic

financing has grown in popularity since the introduction of the Basel III capital rules, which slapped hefty leverage and funding charges on traditional securities financing. Banks found they could lower their capital and liquidity costs by providing leverage via swaps, and then netting and hedging these exposures in the trading book, rather than lending clients the cash to purchase securities outright.

Some clients also prefer synthetic financing because it allows them to sidestep disclosure rules and leverage limits that apply when trading physical securities on margin. Rules introduced by the SEC in 2008 set a 15% floor on margin for portfolios of single stocks, limiting investors to six times leverage.

Margin requirements for total return swaps can be as low as 5%, allowing up to 20 times leverage. “It will depend on the type of securities, but if we’re talking

“

“If we’re talking about developed market, highly liquid, large cap stocks, a 5% initial margin requirement is not that unusual.”

**A lawyer specializing in hedge fund agreements**

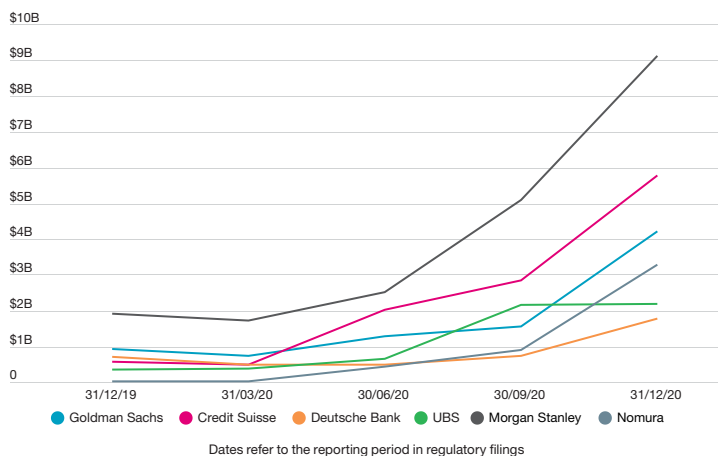
about developed market, highly liquid, large cap stocks, a 5% initial margin requirement is not that unusual,” says a lawyer specializing in hedge fund agreements.

Archegos was required to post just 10% margin against the equity swaps it traded with Credit Suisse, *Risk.net* revealed on April 30.

By using swaps, Archegos was also able to avoid the SEC’s disclosure rules, which require investors to report their holdings if they acquire more than 5% of a company’s shares. Those rules did not apply to Archegos because it amassed its stakes using derivatives, rather than purchasing the shares directly. That allowed the fund to keep its positions largely hidden, even from its prime brokers, who seemed to be caught off guard by the scale of Archegos’ holdings when it defaulted.

To understand how large and concentrated Archegos’s positions were, *Risk.net* analyzed the regulatory disclosures of its

**1. All bank exposures to Archegos-linked stocks over time (\$)**



Source: Regulatory filings and *Risk.net* analysis

six main prime brokers, which bought the stocks as hedges against the derivatives trades. The banks' 13F filings show their holdings of eight core stocks linked to Archegos increased from \$3.8 billion at the end of the first quarter of 2020 to \$7.4 billion on June 30. By year-end, this had jumped to \$26.4 billion (*see figure 1*).

Between them, Archegos's prime brokers owned nearly 50% of the shares outstanding of GSX Techedu, a Chinese e-learning company (*see table A*). They also collectively held a 32% stake in Chinese online video platform iQIYI and over 20% of the shares in Discovery, Tencent Music and ViacomCBS.

Assuming the banks' year-end holdings remained constant, these positions would have been worth a combined \$44.8 billion on March 22—the eve of Archegos's default.

The true figure is likely to have been even higher. Some banks raised their holdings of Archegos-related stocks in the first quarter. Goldman's 13G filings, which disclose trading by parties holding with over 5% of a company's equity, show its stake in GSX Techedu increased from 16.5% to 24.6% on January 29.

Archegos's portfolio began to unravel after shares of ViacomCBS fell 30% between March 22 and 24. When the fund couldn't meet margin calls from its prime brokers, they terminated its swaps and began offloading the stocks they bought as hedges, incurring losses of over \$10 billion in the process.

The spread of losses reveals the good, the bad and the ugly of synthetic prime brokerage.

Credit Suisse reported a \$4.7 billion first-quarter hit from its dealings with Archegos, with an additional charge of \$650 million expected in the second quarter, taking the total loss to \$5.35 billion.

Nomura's losses are similar in scale to Credit Suisse's, relative to its exposure. The Japanese bank took a \$2.3 billion hit in the first quarter, and expects a further loss of some \$570 million from positions still on its book at the end of March.

At the other end of spectrum, Deutsche Bank and Goldman Sachs held more than 20% margin against their exposures, enabling them to exit the trades without

**“It's not that prime brokerage is inherently risky. It's that some of these banks are risky because they are Frankenstein creations and nobody knows what's going on and where the risk is.” Carson Block, Muddy Waters Capital**

reporting losses. Deutsche Bank even returned some collateral to Archegos after winding down its hedges, while Goldman made a small profit, according to a person familiar with the matter.

In the middle ground, Morgan Stanley escaped with a \$911 million loss, while UBS disclosed a \$774 million first-quarter loss.

### Revolving door

Sources say Credit Suisse may have stumbled badly in managing Archegos's default in part because bouts of cost-cutting and a revolving door of leadership left the division rudderless.

The Swiss bank rose to the upper tier of the prime services ranks under Phil Vasani, who led the unit from 2003 to 2013. He staked a bet on the “multi-prime” model, joining forces with technology vendor Paladyne Systems in 2006 to build an open architecture platform that made it easier for hedge funds to connect to several prime brokers.

The multi-prime model gathered steam after the collapse of Lehman Brothers, which trapped billions of dollars of hedge fund assets in years of bankruptcy proceedings. As clients of other US broker-dealers including Goldman Sachs and Morgan Stanley

looked to diversify their counterparty risk, Credit Suisse emerged as one of the big winners.

“Our phone rang off the hook,” says one former Credit Suisse prime brokerage executive. “We got all the business we could handle and we had to turn people away. Credit Suisse was seen as a really good platform. They had built it up for many years prior to 2008 even.”

The Swiss bank peeled away clients—and top executives—from rivals such as Goldman Sachs and Merrill Lynch in the wake of the financial crisis. In 2010, data provider Hedge Fund Intelligence ranked it as the world's second largest prime broker, behind Goldman Sachs.

After turning the prime brokerage unit into one of Credit Suisse's largest revenue drivers, Vasani was promoted to lead the firm's US private banking operations in 2013. He left to join BlackRock in 2016.

Since then, six different prime brokerage heads have led the division.

Vasani was followed in 2013 by Paul Germain, a former Goldman Sachs partner who joined Credit Suisse in 2009. He lasted two years before Mike Paliotta took the reins in 2015. A year later, Indrajit Bardhan was named head of prime services. Bardhan, the former global head of risk for prime services, held the role until 2018, when he was replaced by Paul Galletto. His appointment coincided with the departure of 20 prime brokerage executives, including five managing directors.

Galletto was promoted to run equities sales and trading in 2019. That year, John Dabbs, former head of derivatives prime services, and Ryan Nelson, who ran US financing, were named co-heads of prime services.

**A. Bank holdings of Archegos-linked stocks as of 31/12/20 (%)**

	Baidu	Discovery	Farfetch	GSX Techedu	iQIYI	Tencent	ViacomCBS	Vipshop
Credit Suisse	2.0	9.0	0.8	7.9	6.7	6.4	6.0	3.1
Deutsche Bank	0.7	3.4	2.2	0.1	1.1	1.4	1.0	1.4
Goldman Sachs	1.3	2.1	2.2	16.2	6.6	3.0	1.7	3.0
Morgan Stanley	1.9	6.5	12.8	10.1	8.4	5.8	7.2	7.3
Nomura	0.6	4.5	0.0	7.3	4.2	4.0	3.4	2.7
UBS	0.7	1.7	0.2	7.8	5.3	1.8	1.7	0.2
<b>Total % holding</b>	<b>7.2</b>	<b>27.2</b>	<b>18.2</b>	<b>49.4</b>	<b>32.3</b>	<b>22.4</b>	<b>21.0</b>	<b>17.7</b>

Source: Bloomberg, 13F filings



Then, Credit Suisse lost its global head of prime services risk following the tragic death of Jason Varnish in a ski lift accident in February 2020.

The subsequent appointment of Parshu Shah, a 20-year Credit Suisse veteran, as head of prime services risk has raised eyebrows. Prior to being handed responsibility for monitoring risk in the unit, Shah worked on the swaps sales desk, where he was reportedly in charge of the Archegos account.

Galietto, Dabbs, Nelson and Shah all stepped down following the Archegos blow-up, adding to a list of senior departures that includes group chief risk and compliance officer Lara Warner and the head of the investment bank, Brian Chin.

Bardhan is now returning to Credit Suisse as a consultant to help clean up the mess, according to a Bloomberg report.

### Static charge

Others say this turnover was only part of the problem. At many banks, outdated and cobbled-together risk systems can make it difficult to effectively monitor client activity and share information within the firm.

“Goldman is the one exception where information just flows smoothly,” says Carson Block, founder of Muddy Waters Capital. “It’s not that prime brokerage is inherently risky. It’s that some of these banks are risky because they are Frankenstein creations and nobody knows what’s going on and where the risk is.”

Credit Suisse may have been especially vulnerable to this kind of problem. The bank has a fragmented operating structure—the cash prime brokerage business sits within its US broker-dealer entity, Credit Suisse Securities USA; equity swaps are typically traded out of its London entity, Credit Suisse International; and equity options are generally handled by a separate unit, Credit Suisse First Boston.

Prime brokers use a variety of internal models to set margin requirements. These range from static or rules-based approaches to dynamic, scenario-based

versions. To maximize financing revenues, prime desks are sometimes guilty of choosing the methodology that affords the maximum amount of leverage for a client, sources say.

“That’s what clients want—more margin flexibility,” says the second former prime brokerage head. “Unless a risk manager sees a problem with it, if one model shows a better result for the client, that will be the official one for that portfolio.”

“**You try to come up with something transparent and replicable, yet dynamic. Sometimes clients will pay more collateral if it’s not a black box.” Head of prime brokerage at a large US bank**

Credit Suisse is reported to have used a “static” model to margin the Archegos portfolio. Under this approach, which is fairly common for equity swaps, margin requirements are set at a certain level, subject to predefined constraints, or rules. For example, if a client’s portfolio contains a minimum number of names and the position sizes do not exceed a certain percentage of average daily volume, then the margin charge might be fixed at 15%.

“A rules-based margin calculator is basically a set of constraints that will allow you to land in a certain place,” says the second former prime brokerage head. “Those are not unusual. They are just less sophisticated.”

Sources say the 10% margin charge Credit Suisse applied to the Archegos portfolio was well below industry standards for the level of risk the fund was running. “Given the leverage they had, a really small move in the underlying could create a 10% loss in a heartbeat,” says a senior risk manager at a buy-side firm.

Credit Suisse was reported to be on the cusp of switching to a dynamic model when Archegos defaulted.

The head of risk at a large clearing bank describes static margin as “very 1980s.” Most banks use dynamic, risk- or scenario-based models to margin secu-

rities portfolios and cleared derivatives. These are also standard at clearinghouses. With dynamic margining, the amount of collateral depends on a variety of parameters, including asset prices, volatility and correlation. Additional charges can also be applied to account for concentration, liquidity and basis risks.

The head of risk says his bank uses a dynamic model with a volatility scan range—an estimate of the maximum up or down move in prices—of seven times the daily standard deviation when margining traditional cash prime brokerage portfolios, with a 15% floor for single stocks.

A senior risk manager at another large US clearing bank says prime brokers that use dynamic models will typically collect enough margin to cover a minimum move of 12%, with most applying a price shock closer to 20%.

But convincing hedge funds to accept dynamic margin models, which can seem like a black box, can be tricky. “You try to come up with something transparent and replicable, yet dynamic. Sometimes clients will pay more collateral if it’s not a black box,” says the current head of prime brokerage at a large US bank. “Part of the business is being able to negotiate effectively with clients to ensure you are allowed to change margin multipliers, and change the margin methodology, if it is not working with the level of risk.”

It’s not only on margin models and amounts where prime brokers can low-ball to win business.

In synthetic prime brokerage, swaps are typically margined under the International Swaps and Derivatives Association’s (Isda’s) credit support annex. Under the 2002 Isda CSA—the latest version of the documents—clients typically have one business day to deliver margin before they are deemed to be in default and positions can be liquidated.

Most banks default to this standard timeframe, though lawyers say terms can vary from a matter of hours to as long as three days—the grace period under an old 1992 version of Isda’s CSA.

“When we’re negotiating those financing arrangements, the amount of attention that is given to what time period would there be between the default and the time when the bank



Goldman Sachs is said to be “notorious” for its tough stance in negotiations.

providing the financing would be able to foreclose and liquidate the collateral is immense,” says Fabien Carruzzo, head of the derivatives and structured products group at the law firm Kramer Levin. “Sometimes we’re negotiating by the hour regarding how long a client would have to cure the issue, before the bank can exercise its right to foreclose on the collateral, liquidate the hedges and seize the cash margin.”

According to another industry source, Goldman Sachs is “notorious” for its tough stance in negotiations. “It’s like pulling teeth to try and get any leniency out of them, and they leave a lot of possibilities to close out quickly,” the source says. Other prime brokers can be far more obliging, this person says.

Prime brokers also need some measure of transparency from clients to calculate concentration add-ons, which are bolted on to base margin requirements. But hedge funds can be secretive about their strategies, and banks that push for more information risk losing business.

A senior risk manager at a third large bank says his firm has “lost a lot of clients” due to strict disclosure requirements.

The prime brokerage head at the large US bank says some hedge funds can be downright misleading. “We have seen

clients in the past saying they’re executing a particular strategy and then all of a sudden, out of nowhere, they’re doing something unusual,” this person says. “You need to be able to catch that and understand what’s going on.”

According to a report in the *Financial Times*, some of Archegos’s prime brokers are investigating whether they were “fraudulently induced” to do business with the family office.

Regulators are seeking to bring some kind of order to this corner of the market. Non-cleared equity swaps traded with SEC-registered broker-dealers will be subject to a minimum 15% initial margin requirement under rules that take effect on October 6. Similar rules for prudentially regulated firms will be expanded to hundreds of buy-side clients on September 1. Those rules came too late to contain Archegos, and many hedge funds may still escape the regime. Firms with less than \$8 billion of non-cleared swaps are exempt from prudential versions of the rules, while a \$50 million exchange threshold, included in both iterations, means hedge funds that spread their business across multiple prime brokers may be able to avoid posting margin under the new rules, even if they are in scope.

The SEC is also working on new reporting rules for investors that build stakes in companies using total return swaps.

Credit Suisse has already pledged to cut leverage exposure in prime brokerage business by \$35 billion, or a third.

Others are also taking steps to reduce risk. “We have been having conversations with our clients and telling them when their positions are too risky for what they’re paying us and that they have to think about things differently,” says the current head of prime brokerage at the large US bank. “This will be a continuing theme across the industry.”

The lawyer specializing in hedge fund agreements expects to see prime brokers ditching the use of static margining for equity swaps in favor of dynamic models that can be used to cross-margin cash and synthetic exposures. “Prime brokers are moving toward cross-margining, having swaps margined in this more cash prime brokerage way,” he says. “I think this is only going to increase, because if you’ve locked margin in at 5% and you can’t change it, then that might really hinder your ability to react.” **wt**

*Additional reporting by Will Hadfield.*



# Despite client concerns, LSE's new Sedol fees take hold

Under the new fee policy, some of the largest users of the LSE's identifier codes could see their Sedol spend more than double, though the exchange says the 'vast majority' of clients will see no increase. By Max Bowie



Five months into the rollout of a new pricing policy for its Sedol global identifier code, some of the London Stock Exchange's data clients have not yet signed a declaration acknowledging the new fees. And while LSE officials say the number of holdouts is dwindling, some client firms—and the industry associations that represent them—continue to voice objections to the new policy.

The LSE says the changes reflect a significant expansion of Sedol codes over the past decade, and a more equitable fee structure, which in the past did not fairly represent usage across the spectrum of clients, especially the largest firms.

James Nevin, managing director and head of fund and research products at the LSE, says that over the last decade, the exchange has expanded the number of Sedols from 1.2 million, which were largely focused on UK securities, to a global, multi-asset identifier numbering more than 110 million, of which 20 million are live securities.

In addition, within the last year, the exchange has added API access to Sedol codes, and has increased the refresh rate from three times per day to hourly, and extended Sedol codes to cryptocurrencies and other digital assets.

However, user groups that represent LSE clients among buy-side and sell-side firms say the increases are “unreasonable and unfair,” as the exchange is increasing costs without adding value.

The key policy change moves away from assessing fees based on the number of legal entities within a firm that consumes Sedol codes, to a model based on the number of business segments that consume Sedols in different regions.

For those with “limited” user licenses (i.e., those using fewer than 10,001 Sedol codes), fees will not change. A firm in Band A (which uses fewer than 1,000 Sedol codes within one legal entity) would pay £860 (\$1,200) per year. A firm in Band B (1,001 to 5,000 Sedols) would pay £2,715. And a firm in Band C (5,001 to 10,000 Sedols) would pay £8,925. Distribution licenses (which allow firms to distribute Sedol data to licensed third parties) for the same bands would be £795, £2,475, and £13,315, respectively.

The fees resulting from the policy change will be phased in over two years, starting from January this year, with 50% of the increase taking effect in 2021, and the remainder in 2022.

During the “transition year” of 2021, the LSE will calculate fees based on a combination of its old legal entity-based model and the new business segment and region-based model. So during 2021, firms with unlimited user licenses (more than 10,001 Sedol codes) will be covered by three further bands of usage and fees: Band D (more than 10,001 Sedols used by one legal entity) will cost £29,070. Band E (more than 10,001 Sedols used by up to five legal entities) will cost £91,865. And Band F (more than 10,001 Sedols used by six or more legal entities) will cost £121,310.

When the new policy takes full effect next year, firms will pay fees ranging from £35,000 for usage of more than 10,000 Sedols by one business segment in a single region up to £251,000 for usage by more than five business units across more than three regions (*see fee matrix, below*).

Full details of the policy are available on the LSE's website, along with an online fee calculator to help firms understand how much they should pay.

Between 700 and 800 client firms will be affected by the policy, Nevin says. The vast majority of clients will either see no fee increase or will actually see

“

“We agree it's reasonable to pay some kind of increase, but not that much.”

Data executive at a large European bank

their Sedol costs fall as a result of the new policy, but he acknowledges that larger firms will see costs rise—in some cases, significantly.

“Around 30% of that affected part of the client base have seen a price decrease, and around 45% are fairly flat or have seen an increase of less than £10,000. Obviously, that leaves about 25% whose fees have gone up—some significantly,” Nevin says. “Generally, the rebasing of the policy better aligns with the differentiated usage across our customer base. We had a lot of mid-range clients who were previously at the same level as big banks, and those levels now more appropriately reflect usage across the client base.”

The policy also offers degrees of flexibility: For example, if a firm has a global equity research arm based in the UK that would fall in the 10,000+ Sedol usage tier, but also has a US business that only trades US equities and would therefore use fewer Sedols, the firm's contract could reflect those different tiers of usage.

“Our sales team helps maximize the policy for clients' benefit. We don't

## Fee Structure: Unlimited User Licence Fee Levels



The below fees will apply from 2022 for Unlimited User Licence customers (>10,000 codes) under the new Business Segments & Region model:

2022 SEDOL Unlimited User Licence Fees (per annum)

		Regions		
		1	2	3+
Business Segments	1	£35,500	£66,000	£87,000
	2	£66,000	£122,000	£161,000
	3	£87,000	£161,000	£212,000
	4	£97,000	£180,000	£238,000
	5+	£103,000	£190,000	£251,000

LSE Group's new Sedol fees for 2022 onwards (source: London Stock Exchange Group)

just automatically throw them into the highest tier,” Nevin says.

Indeed, one market data manager at a brokerage falling within the lower tiers confirms that his firm is not impacted

“Based on market data user groups’ assessment, the vast majority of members are significantly impacted by the new licensing model imposed by the LSE on Sedol identifiers. In other words, all or most sizes of firms are impacted and see significant increases of their Sedol licenses.” Cossiom executive

by the change, but also notes that the firm deliberately “barely uses Sedol” specifically because of the cost.

However, a data executive at a large European bank says its Sedol costs will increase by 200% under the new policy, even though the bank’s use of the codes is not changing. “We agree it’s reasonable to pay some kind of increase, but not that much,” says the source.

### ‘We’re trying to treat the overall market fairly’

The LSE denies it is using a monopoly position to wring more fees out of customers. First, it says the overall increase is not so big, and second, it says there are alternatives to Sedol. “Overall, yes, we are making money from the policy change, but not as much as clients may think,” Nevin says. Under the old policy, prices grew at a 3% compound annual growth rate (CAGR) between 2010 and 2020, while the new policy would see CAGR for the period 2010 to 2022 as the policy takes effect rise to 3.9%, he says.

“Most of the feedback has been positive because most firms’ fees have gone down or remained flat. I’m not going to say everyone is happy about it ... but even though the larger-tier firms don’t like increases, most understand why we’re doing this, and that we’re trying to treat the overall market fairly,” he says.

He says the majority of clients—though not all—had signed a declaration for the new policy by the

January 1, 2021, effective date, and the number outstanding is becoming smaller as time goes on. Nevin says the LSE has approached this by working with firms on a client-by-client basis to understand their individual pain points and reasons for delaying.

“For example, some are not pushing back on the policy, per se, but ... because it is widely used, they need to analyze their internal use and do due diligence,” he says.

In an attempt to minimize these delays to adoption and make the transition as smooth as possible, the LSE engaged with UK-based industry user group Ipug and French user association Cossiom to conduct surveys of their memberships. This follows an attempt to revamp the Sedol policy in 2018, which the LSE ultimately abandoned after significant pushback from clients, deciding instead to try to engage the industry in creating the new policy.

“In 2018, we looked to do a fairly crude tweak to our commercial policy ... based on how many legal entities used Sedol within an organization. And we got some decent pushback. Clients said legal entities weren’t a good measure of usage,” Nevin says. “So at that point, we went into a period of consultation, including more than 40 meetings with clients and industry bodies. That concluded at the end of 2019, and we announced the new policy early in 2020, with implementation starting in January 2021, phasing in over two years.”

However, though Nevin says feedback has been mostly positive, the user groups aren’t satisfied with the results, and take issue with the LSE’s assertion that only the largest firms will be affected by the changes, saying that the multiple dimensions of the licensing model make it hard to optimize usage, and therefore tacitly encourage firms to sign up for higher-tier licenses.

“Based on market data user groups’ assessment, the vast majority of members are significantly impacted by the new licensing model imposed by the LSE on Sedol identifiers,” says a Cossiom executive. “In other words, all or most sizes of firms are impacted and see significant increases of their Sedol licenses. We don’t acknowledge that

only big firms are impacted,” adding that in a letter to the LSE last year objecting to the changes, Ipug stated that 80% of members reported “unacceptable cost increases of greater than 80%, and one in three firms reported increases of more than 100%.”

The user groups are also unhappy that the LSE began its consultation by working with them to survey members in a constructive way, but then they say the exchange did not incorporate their feedback, and declined to work further with them, instead negotiating individually with firms.

Nevin says user groups proved very helpful in terms of translating feedback from their members, but adds that once the policy change took effect at the start of this year, the LSE felt that any further dialogue should be conducted with firms on an individual level.

### ‘Resistance, resentment, but resignation’

The user groups continue to object to the new policy, and advise members against signing it, but ultimately, they have few options. If any firm refuses to sign the declaration, LSE can turn off that firm’s supply of Sedols, by asking its vendors to switch off the codes to that client. “That’s not easy, because most big firms take Sedols via multiple vendors. But I don’t think it will come to that,” Nevin says, adding that the exchange has not lost any Sedol clients yet as a result of the new policy.

But clients say Sedol codes are a necessity, even if their UK trading activities are relatively small. “We cannot just stop using Sedol codes; we need those codes in our systems. If we don’t have the codes, we can’t trade,” says the European bank data executive.

The Cossiom exec concurs, adding: “We believe that a firm does not have the option to use Sedol or not. Sedol identifiers are embedded into too many information systems and too many workflows. They are indispensable to ensure interoperability across the financial industry and for clearing on UK markets. A firm cannot decide not to use Sedol identifiers.”

Given this industry dependence on Sedols, particularly for trading on UK



**James Nevin**  
LSE



markets, “It might be very difficult or even impossible for most financial institutions (notably tier-one and tier-two buy-side and sell-side firms, and securities services firms) not to use Sedol identifiers anymore,” the Cossiom executive adds.

Nevin adds that while still best-known for their uses supporting trading and post-trade clearing and settlement, Sedols are used across firms’ entire investment lifecycles, including for valuation.

“The value is not just the code itself, but that it validates an amount of fundamental data that it represents, and that you can use it to map that web of underlying fundamental data. So the code is

not just one data point—it’s that mapped data so you can match trades and map to research data,” Nevin says.

Among end users, there is “resistance, resentment, but ultimately resignation” toward the Sedol policy, says one industry consultant familiar with the situation, adding that this reflects a worrying trend of data providers increasing fees but engaging less with the clients who are subject to those fees. Ipug, for example, is developing a scorecard-based rating for all suppliers that the user group hopes will better inform consumers and regulators, and will positively influence suppliers’ policies and encourage more engagement.

The Cossiom executive also says user groups must remain important for collectively representing their members in negotiations where they believe policies have industry-wide impact.

“We believe it is our duty to take the right time and share our concerns with vendors when it is believed that a new commercial model or initiative is not relevant or reasonable,” says the executive. “When most market data user groups across Europe believe that a new commercial model is not reasonable, it should not be unrealistic that the vendor listens to clients, reviews its model, and comes back with something acceptable. This is what user groups expect to achieve.” [WT](#)



# Swipe left: repo reporting no match for Brexit, or collateral



While it enjoyed a happier start than Emir, the SFTR's honeymoon is over now that trades report separately in the UK and the EU. By [Samuel Wilkes](#)

**A**s ill-fated regulatory encounters go, few can rival the star-crossed introduction of derivatives reporting in Europe. Repo and securities lending participants have taken comfort in the fact that their own entrée to the Securities Financing Transactions Regulation (SFTR), which governs repo reporting, was not quite as disastrous.

Even so, reconciling counterparty reports for the same repo trade—known as pairing and matching—has troubled the SFTR's debut, sources tell *Risk.net*. A list of issues maintained by the International Capital Markets Association (Icma) is said to contain more than 50 separate items that either cause rejections in reports or

breaks in reconciliation when pairing and matching trades.

And while reports are failing to reconcile for a whole host of reasons, the blame has been squarely placed on two main culprits: Brexit and collateral. The first is beyond counterparties' and trade repositories' control—but the second has firms remediating errors and seeking clarification from regulators.

"Pairing and matching are still much higher than expected, especially compared to the relevant precedents we've seen," says Alexander Westphal, a director in market practice and regulatory policy

at Icma. "But clearly there are still significant issues—especially on the collateral side—that need to be sorted out."

The head of regulatory reporting at a European asset management firm agrees: "[Collateral would be a] good topic to get some light on, as this is burdensome for doing the reporting."

Sources say pairing and matching rates took a knock following the UK's exit from the European Union. Jonathan Lee, a senior regulatory reporting specialist at technology provider Kaizen Reporting, estimates that, based on public data released by repositories, pairing rates





dropped from between 40% to 50% to mid-20% between November 2020 and February 2021. A senior compliance manager at a European bank concurs with those figures.

*Risk.net* submitted a freedom of information request to the European Securities and Markets Authority (Esma) to verify these figures, which the watchdog denied on the grounds that it does not maintain the figures in an internal document. Esma has granted a *Risk.net* FOI request for the derivatives reporting regime's pairing and matching in the past.

“I believe matching will receive more attention after summer, once SFTR has been live for a year.” **Jesper Lerche, Arbejdernes Landsbank**

“We have seen some firms, even to this day, have good pairing rates at 90% paired, but that doesn't seem to be standard across the board. Pairing and matching rates are still very low,” says Joanne Salkeld, SFTR product manager at the reporting arm of MarketAxess. “I think throw

Brexit into the mix and it has caused additional trouble for pairing and matching.”

### Every single day

Reporting under the SFTR requires parties to a repo or securities lending transaction to submit information on the trade to repositories before the end of the following working day. Repositories must then reconcile critical fields within reports of the same trade in the process. For banks, central counterparties and non-bank firms regulated under Mifid II, this began in July last year. Fund managers not authorized under Mifid II began reporting in October 2020 and corporates in January 2021.

Pairing is the first step, and requires repositories to ensure reports of a single trade have the same unique code identifying the specific trade (UTI), the same type of governing master agreement, and that the reporting parties have listed each other as counterparties in their reports.

Matching is the second step, once reports have been paired. Reconciliation requires more than pairing for a successful match—reports must match information in 62 fields describing key attributes of the trade.

Dual-sided reporting allows regulators to verify the information they receive is correct. Reports that fail to successfully reconcile indicate to regulators something is inaccurate in the reports, which repositories and counterparties must figure out and remediate.

Many acknowledge that pairing and matching rates are higher than they were at the outset of derivatives reporting under the European Market Infrastructure Regulation (Emir) (see box: *Learning from Emir*). But others strongly disagree. One source even describes the rates as “really grim.”

Jesper Lerche, a senior business analyst at Danish bank Arbejdernes Landsbank, says failure of reports to pair or match happens every trading day. He estimates it typically takes one to two weeks to resolve with counterparties, but the worst cases can take as long as a month.

And not everyone is trying to resolve breaks in pairing and matching. The senior compliance manager at the European bank says it is focusing on



**“It was a surprise that master agreement type, along with UTI and LEI, should be part of the pairing, and frankly somewhat impractical.” Jesper Lerche, Arbejdernes Landsbank**

remediating issues with pairing for now and will turn to improving its matching statistics in the future.

“A lot of entities have had trouble with stabilizing the whole SFTR setup, so I think some banks have not had much spare energy to deal with matching issues,” says Lerche. “I believe matching will receive more attention after summer, once SFTR has been live for a year.”

### Brexit breakdown

A major factor in Brexit's effect on the decline in pairing is the increase in the number of single-sided reports being submitted separately to EU and UK trade repositories (TRs). Post-Brexit, EU and UK regulators no longer have access to reports submitted by counterparties in one another's jurisdictions, which means trade repositories keep reports of EU and UK firms separate.

But because cross-border trading between EU and UK firms still occurs, trade repositories are unable to reconcile their reports. This is good news for banks, as it means they have finally been granted their wish for single-sided reporting.

“Now, all of a sudden, you're not sending your UK transactions to EU TRs but the EU entity you are trading with is still sending its EU transactions to the EU TR,” says the senior compliance manager at the European bank. “Those reports are single-sided now and they're not going to pair with anything—which in effect is not a bad thing, because then you have less to worry about—but the noise is going to cause pairing rates to skew toward the bottom.”

The fall in pairing rates will be particularly high for reports where buy-side firms in one jurisdiction had been delegating their reporting to their dealers in the other—the details of these were more likely to reconcile as they are submitted by the same entity. But as dealers now only need to report their side of the trade in their jurisdiction, where those trading relationships still exist, there is no longer a guaranteed match but a guaranteed failure.

“Where you have delegated reporting, I would hope they would both pair and match on the basis that there is only one party responsible for reporting—that was where the strength in the rates was,” says Lee of Kaizen Reporting. “Now both sides aren't reporting to one single regime, so those pairing and matching rates have fallen because of that. Additionally, there is simply a great deal more single-sided reporting.”

Although repositories are unable to reconcile these reports and don't attempt to do so, they still count toward pairing and matching rates.

“The pairing rates are literally a yes-or-no binary calculation,” says Nicholas Bruce, head of business development at Regis-TR, a trade repository. “So, while in practice we only pair eligible trades, excluding one-sided trades from the process, the overall pairing rate is typically calculated based on the total universe of trades reported. Clearly the rate would be very different if you only looked at the trades available to pair.”

Branches of EU firms in the UK have also caused an increase in the number of single-sided reports, as they are required to submit reports to both the UK and EU repositories for trades they execute.

Because they are reporting trades to repositories on both sides of the channel, one side will always have a single-sided report. If the UK branch of an EU dealer trades with an EU firm, UK regulators will receive only the report from the UK branch—not from the EU client. Correspondingly, EU regulators will see only one side of the report when the UK branch of an EU dealer transacts with a UK client.

Trades with branches also cause breaks in matching rates, as their counterparties don't always identify that the dealer is acting through its branch in the fields describing the counterparty.

### Collateral damage

Brexit aside, many different things prevent reports being fully reconciled—the high number of fields that needs to be paired and matched increases the chances of a break. According to the senior compliance manager at the European bank: “It could be anything.”

One issue, says Arbejdernes Landsbank's Lerche, is including the type of master agreement that governs the trade.

“It was a surprise that master agreement type, along with UTI and legal entity identifier (LEI), should be part of the pairing, and frankly somewhat impractical, since master agreement type is a field where there could easily be matching differences,” says Lerche.

Meanwhile, far more issues cause breaks in matching than in pairing due to the high number of fields that must be reconciled, which will increase from 62 to 96 fields from April 2022.

### Learning from Emir

Legislators and regulators learned their lessons from the rocky start of Emir reporting and safeguarded against the same issues occurring when SFTR reporting began.

One such lesson was to avoid a logjam of counterparties trying to onboard with all trade repositories at the same time just a month before the start of reporting. Under Emir, repositories were hit with an avalanche of onboarding requests overwhelming their systems and causing backlogs in reports.

For the SFTR, legislators set out a staggered start, as noted above, in which different types of market participant began reporting at different start dates.

“Everyone was generally well-prepared, having learned the lessons from the launch of Emir,” says Nicholas Bruce of Regis-TR. “Also, the tranches rollout led to a more measured approach across the market and allowed repositories to handle the flows better. There was also more testing by market participants and intermediaries before reporting began.”

A further lesson Emir provided was to assign responsibility for the generation of UTIs, which allow regulators to identify separate reports of the same trade. At the start of Emir, there was confusion over which parties should generate UTIs, leading to counterparties reporting separate UTIs for the same trade.

Firms must either agree bilaterally who generates the UTI or follow a sequence of actions that assigns responsibility to different parties to the trade in different circumstances—in a so-called regulatory waterfall.

For cleared trades, the clearing member generates the UTI for its trade with a client and the central counterparty generates the UTI for the member to transfer the trade for clearing. For non-cleared trades executed on-venue, the trading venue generates the UTI and for off-venue centrally confirmed trades, a confirmation platform is used. Where trades are directly executed between financial counterparties, the collateral provider generates the UTI in the case of a securities lending or borrowing transaction and the collateral receiver in the case of a repo transaction. If one of the counterparties is non-financial, then the financial counterparty is responsible for the UTI.



Two separate sources—Kaizen’s Lee and the senior compliance manager at the European bank—say matching rates are as low as 10% to 11%.

A wide range of sources point to the matching fields for collateral being a big culprit behind reports failing to match.

“If you look at the collateral reconciliation results, there are many fewer fully reconciled reports than the loan reconciliation,” says Catherine Talks, SFTR product manager at trade repository Unavista.

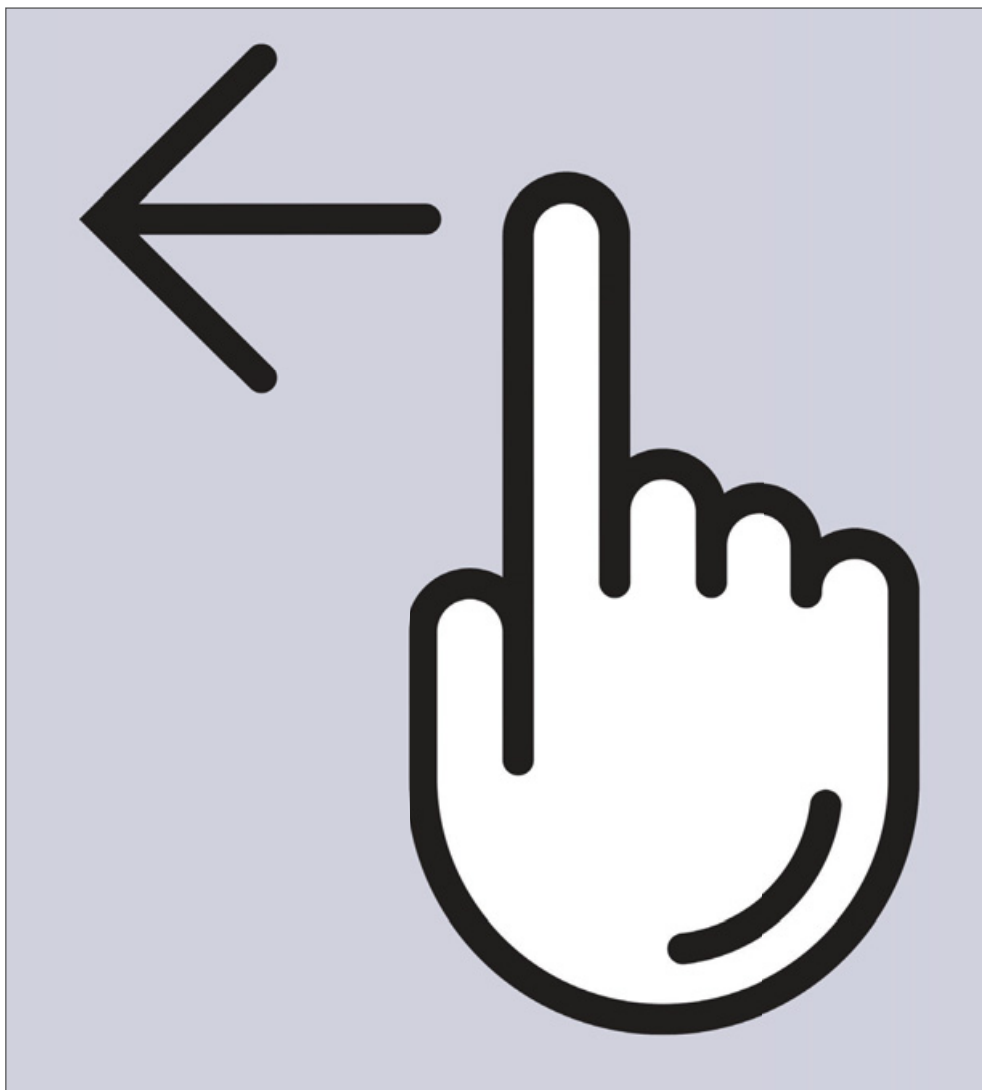
Sometimes, the breaks are caused at the level of the trade repository. One issue is the messaging format that trade repositories use to communicate updates to reports with each other.

Two repository sources say the current messaging standard doesn’t recognize negative values, which it needs to do for circumstances in which a counterparty is giving collateral. Instead, trade repositories submitting reports in which both parties show positive values cannot match—the report of the collateral giver needs to be negative.

“When reports are sent through the inter-TR process, there is a schema problem for repo and securities sell/buy-backs that means nominal amount will always break,” says Talks. “This is because the rules require a positive to reconcile to a negative value—however the schema does not allow for a negative sign.”

Ian MacKay, product owner for global post-trade services at SFT infrastructure provider EquiLend, says one reason collateral can result in breaks is that firms often use baskets of collateral for the SFTs. As baskets can contain many different securities and be used for many separate transactions, it means there are greater chances of a break.

“The regulator is looking to be able to review the collateral against those trades to monitor market exposure,” says Mackay, but if there is a problem with reporting the collateral basket because the mandatory data fields are not populated, there can be instances where the collateral file would not be processed, he adds. The TR will therefore reject the file and will not be able to match or pair the collateral, resulting in the lack of transparency that the regulator is looking to address.



### Are we a match?

One of the discrepancies the senior European bank compliance manager singles out is in the LEI—the unique code used to identify companies—specifically for the issuer of the collateral used in a repo transaction.

They say they often find mixups with a security issuer’s LEI if it is a multinational with many LEIs corresponding to legal entities in different jurisdictions. That problem should abate over time for European securities due to a partnership between the Global LEI Foundation and the body responsible for generating identification codes for individual securities, known as the Association of National Numbering Agencies.

The collaboration links the LEIs of issuers to the corresponding

International Securities Identification Number (Isin)—a unique identifier given to newly issued securities—which means firms can more easily identify the LEI of the issuer of a particular security.

Most European national numbering agencies have joined the initiative, but not many outside Europe, which means there is no mapping of LEIs and Isins for securities in those jurisdictions.

A further issue affecting both collateral and loan reconciliation occurs among fields for describing the type of collateral or security on loan. There are eight different options to identify collateral received, but some securities fall into multiple categories—such as government bonds that have been repackaged as structured products. [wt](#)

*Additional reporting by Rebekah Tunstead.*

# In fake data, quants see a fix for backtesting



Traditionally quants have learned to pick data apart. Soon they might spend more time making it up. By **Rob Mannix**

It wasn't a surprise that videos of Tom Cruise playing golf and doing magic tricks should rack up millions of views on TikTok earlier this year. The real surprise was that the clips didn't feature Tom Cruise at all. They were fakes.

The technology behind these ultra-realistic "deep fake" videos, now common on social media platforms, has already found a home in finance where quants are using it to create parallel universes of data to test investment strategies.

Experts say this synthetic data could help overcome weaknesses in backtesting, which relies on just a single time series of historical data and says nothing about how strategies

might have fared in different conditions. What quants refer to as synthetic, artificial, or simply fake data, offers a chance to invent alternative histories for deeper testing.

"The data is anchored to real life but altered in various ways," says Joseph Simonian, quant research consultant and co-editor of the *Journal of Financial Data Science*. It gives a "new flavor" to backtesting. A buy-side quant says fake data makes it possible to explore the market's unknown unknowns.

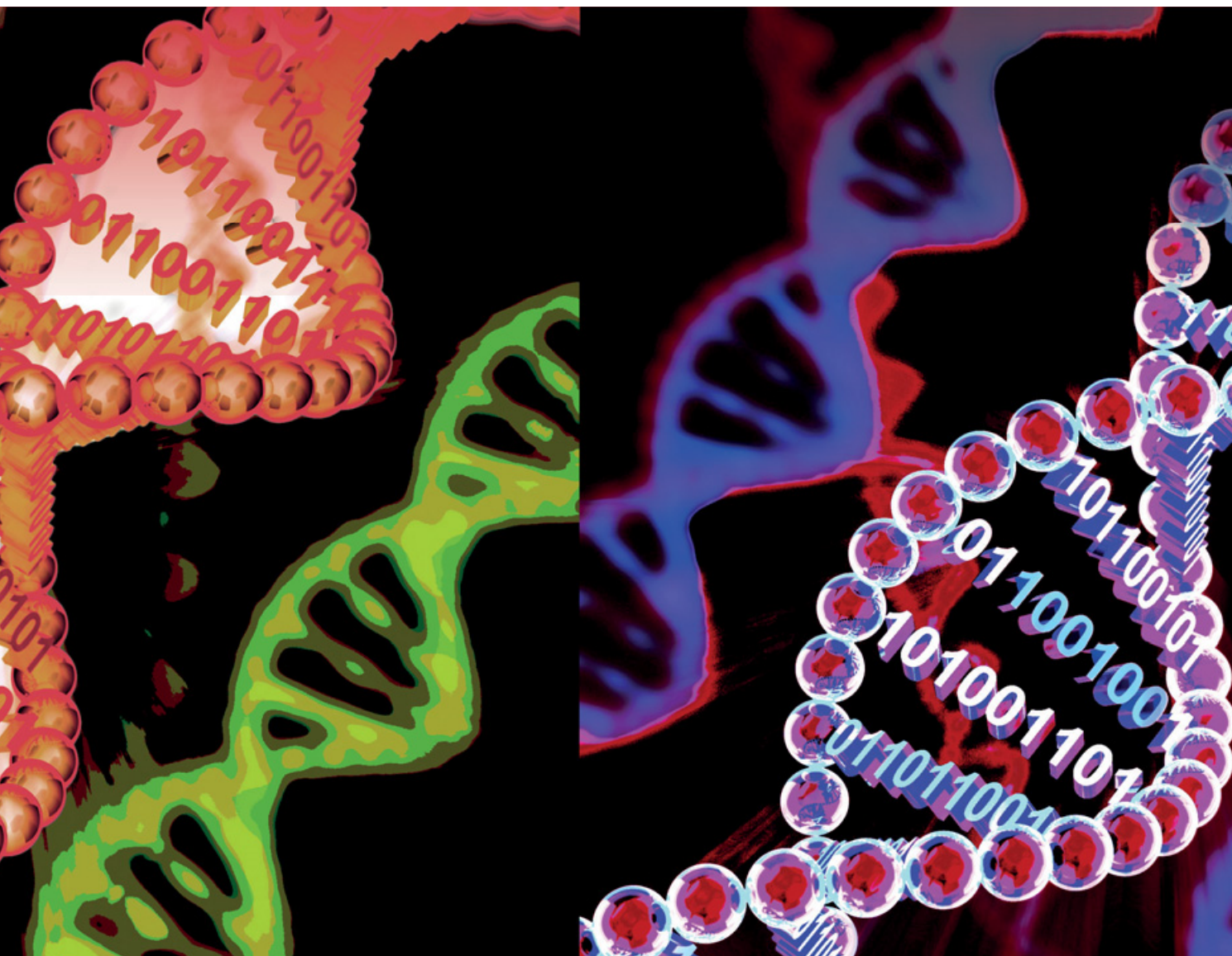
Firms are starting to experiment with the idea and are achieving some notable results. The models can be glitchy, still.

But the technology to apply them in investing is moving fast.

"We can create synthetic series that are indistinguishable from the original," says Blanka Horvath, an academic at King's College London and the Technical University of Munich, and part of a team of researchers and JP Morgan quants working on data generation models. "That's where the excitement is coming from."

Firms of all stripes are trying out fake data. Amundi, Europe's largest asset manager, has begun using synthetic data to test some of its volatility trading and risk parity strategies. ETS Asset





Management Factory, a firm that licenses machine learning algorithms to investors, has used artificial data to test its currency trading algorithms and to develop new models. Deutsche Bank is looking at the new simulation techniques.

And the results demand attention.

Horvath and other academics and the team at JP Morgan have built a data generator for use in training and testing models for options hedging. The models use machine learning to hedge complex derivatives books, in a process known as deep hedging. The data generator produced fake market data that mimicked the original to a 99.9% confidence level.

“We can create synthetic series that are indistinguishable from the original. That’s where the excitement is coming from.”  
**Blanka Horvath, King’s College London**

“These generative models have the potential to approximate time series so much more closely than regular stochastic models,” Horvath says. In fact, there is a theoretical proof that with neural networks quants can replicate a real dataset as closely as they wish. “We can do it as precisely as we want,” she says.

In 2019, a machine learning model built by Magnus Wiese, a quant researcher who splits his time between JP Morgan and the University of Kaiserslautern, together with colleagues bested a Garch model in faking daily S&P 500 prices. The so-called convolutional neural network model uses techniques initially designed for image recognition. Garch models are widely used by quants in simulating market paths.

Bank of America, ETS and another quant firm, Cohen & Steers, separately have trained machine learning models using fake data and tested how successfully the models would invest. With more



data to learn from—and no apparent loss of accuracy—those trained on fake data performed better.

BofA's neural network achieved a higher hit rate compared with training solely on real data. It also scored a better R-squared, which is a measure of how well a model explains the data it applies to. When used to forecast moves in US Treasury bonds over daily to monthly horizons, the model recorded a hit rate as high as 82%.

According to Stefan Jansen, a consultant on the use of machine learning in quant trading and the author of a widely used text on the topic, up to a quarter of the big hedge funds he speaks to already are exploring the use of synthetic data. "It's too promising to pass up," he says.

### Making it up

As far back as the 1980s, trend followers manufactured simple artificial time series such as basic saw-tooth price patterns, to work out how different market conditions would affect the strategies they were developing.

Rudimentary methods of creating artificial data proved especially useful during the last decade in the aftermath of the European sovereign debt crisis, when the European Central Bank set



**Blanka Horvath**  
King's College  
London

cine, scientists tracking Covid vaccine efficacy can access millions of separate time series of data, one for each vaccinated individual. Backtesting can be compared to gauging vaccine efficacy based on a single patient.

So-called bootstrapping and Monte Carlo simulations try to get around the problem. In bootstrapping, quants glue together bits of the past to fashion new versions of history against which they can test ideas. In Monte Carlo simulations, they create plausible future paths for time series based on models of how markets work, adding an element of randomness.

But even Monte Carlo methods follow a data generation process according to a pre-programmed model. This shapes the distribution of the data that the process creates. The convoluted patterns of real markets are impossible to capture perfectly in a mathematical model. "The real-world data is not so neatly described," says Anthony Morris, head of quantitative strategies at Nomura.

Bootstrapping runs into the same barrier, he says. "Mixing up historical data in different orders will distort the nature of actual serial and cross-sectional dependencies. These exist but can be quite difficult to specify," he says.

When markets evolve in ways that make the past a bad proxy for the future, as arguably is the case right now, the problem becomes acute.

### Unknowns

Data generators offer something different. They are able to reproduce complex patterns from real markets, says Thierry Roncalli, Amundi's head of quantitative research: non-linear autocorrelation, fat tails, heteroscedasticity—such as varying cross-sectional correlations—and non-stationarity, the way in which the data's distribution changes over time.

Using generators, quants say they can add richer data into simulations, like how two indexes move together. They can incorporate data from outside the price time series, like sentiment scores or trading volumes.

Synthesizing data, then, can help prepare for unseeable risks—the kind of tail events, like Covid-19, that have no parallel in the historical data.

Amundi is using fake data to test the resilience of its volatility strategies in the face of such events. And in its risk parity strategies, the firm is using the data to calibrate stop-loss and stop-gain mechanisms based on a better grasp of the extremes of the return distribution.

In both cases, testing with synthetic data exposed risks that otherwise would have gone unnoticed, Roncalli says.

"For risk parity, the tail of the probability distribution [of maximum drawdown] ... is less fat than that generated by the bootstrap sampling method but has several extreme severe scenarios," the Amundi team wrote in a paper on their work.

With fake data, quants might also test strategies against known unknowns—periods of stress for which parallels do exist but where the data is scarce or unreliable. An example is a rise in inflation. Markets have enjoyed low inflation for the past 40 years, but many investors in the US and Europe are concerned this may be about to change.

To test an inflation strategy with a bootstrapping method, investors would cherry-pick relevant periods of data—such as stagflation of the 1970s. But slicing data into regime-specific chunks cuts into an already sparse body of information.

It's also unclear that past episodes will necessarily be a good source of information. Previous spells of inflation came before markets globalized, for example.

"You'd like to backtest in periods that are representative of the appropriate market conditions or market regime," said Yigal Jhirad, head of quantitative and derivative strategies at Cohen & Steers, speaking at a recent conference. "[Synthetic data] provides a mechanism with which to develop backtesting models that are regime-specific."

The new models open up a way to a more sophisticated, forward-looking testing, compared with the traditional method of backtesting strategies, which is by definition backward-looking. Quants might input into a model the level of market volatility or features of recent price moves, and ask the model to plot future paths based on those readings, Horvath explains. "We can ask the



**"[Synthetic data] provides a mechanism with which to develop backtesting models that are regime-specific." Yigal Jhirad, Cohen & Steers**

negative rates and quants had to overhaul financial models that were not built for such a possibility.

But traditional backtesting—where investors test the effectiveness of a strategy by charting how it would have performed in real-world conditions—uses only one version of history. "They see only what actually happened," says Jacques Joubert, founder of Hudson & Thames, a firm that licenses quant algorithms.

This creates an in-built limitation for testing. To draw an analogy with medi-



model: If we ended up in a year's time in [a given state of the world], how would the future look after that?"

Elsewhere, Rob Carver, a former portfolio manager at Man AHL and writer on quant investing, says synthetic data could help in stress-testing strategies for which losses can be highly path-dependent.

For trend following and risk parity strategies, markets that slowly drift in the wrong direction can be worse than markets that jump quickly, he points out. (Greater volatility would trigger risk management mechanisms that cause the strategies to reduce leverage.) Simulating market paths and generating a theoretical P&L for such strategies can be more revealing than a point-in-time stress scenario, such as assuming sudden big falls across markets.

And fake data also could help root out strategies that are spurious. Strategies that quants are able to identify in their

research in only one synthesized dataset are probably a fluke, say Gautier Marti, a quant researcher at the Abu Dhabi Investment Authority. Marti is widely considered to be one of the architects of using image generation technology to create synthetic financial data.

To compare the robustness of different investing ideas, Simonian suggests using machine learning iteratively to generate fake time series that inch closer and closer to replicating the real data. By testing strategies on the fake data in each step, quants could measure how far the future would need to diverge from the past before a strategy might break down.

### Trial and error

To be clear, using machine learning to create synthetic financial data is a nascent area. Generative adversarial networks (Gans), one of the models for making fake data that has attracted most interest, were invented only in 2014.

Gans use so-called generator and discriminator neural networks working in pairs and competing with each other through thousands or millions of iterations. The generator tries to learn to create fake data good enough to fool the discriminator. The discriminator effectively tries to call the generator's bluff. Their first use to synthesize time series data came as recently as 2017 and not in finance but to train a model to warn of medical emergencies in a hospital intensive care unit.

Generating and using synthetic data in finance will take time to get right.

Firms have to experiment with different data, decide which relationships they want to model, choose whether to use other variables such as sector or macro information in the generation process. They must pick a model network architecture, set training rates and learning horizons, and select from multiple algorithms for the task. This isn't easy to

determine in advance, Jansen says: “You have to try it out until you stumble on something that works.”

Firms may need to work on the problem “in a concentrated fashion” for six to 12 months, he says. Jansen reckons such a project should be carried out as part of an effort to build deep learning capabilities, which he estimates would require a team of four or five specialists and cost maybe \$2 million to \$3 million a year.

“Gans are really difficult to calibrate. When you change the parameters, Gans are more sensitive. But we can’t say one method is better than another because this is a work in progress.” **Thierry Roncalli, Amundi**

The time required to run the generators holds back the pace at which such research can occur. Running models can take many hours, up to more than a day, practitioners say. And Marti estimates no more than a hundred individuals in Asia and Europe are working in finance who are fluent with the most cutting-edge techniques.

Gans in particular can be infuriatingly hard to calibrate. The models are prone to what’s called mode collapse, in which the model picks up too early on limited features of the data and learns to follow only one possible path without exploring further. “The generator creates data that looks real but you miss lots of possible scenarios,” says Marti.

A neural network with too many nodes effectively learns in too much detail, just like any overparameterized model. This means it might pick up on noise in the training data, like homing in on fuzziness in an image, and learn to generate new data in which the fuzziness dominates.

Quants might calibrate the discriminator poorly. “If you use the wrong optimization objectives you might think your Gan has converged but you produce synthetic data that turns out to be missing the features that really matter,” Horvath says.

In one set of experiments where Jansen scaled up a basic Gan time-series model to synthesize prices and returns for 50 stocks, the fake histories appeared less convincing, losing some of the fat-tailed nature of real market data.

“You may need to overlay these models with some heuristics,” says a buy-side quant. “You might condition the model to make it more prone to come up with appropriate solutions and not to get stuck somewhere in left field. There’s a lot of work you need to do to get these things to output practical results and not just a bunch of gibberish.”

The biggest challenge, though, is simply knowing whether the output data from models can be trusted.

To train a Gan, or to be confident in data from other simpler generative models, quants need a way to determine whether the fake data captures the properties that are relevant from the original. “Honestly, that’s an open problem,” says George Lentzas, chief investment officer at quant firm Springfield Capital Management and an adjunct professor at Columbia Business School.

### Advances

That said, innovations and advances are coming thick and fast.

Academics in 2020 built Stock-Gan, a neural network that creates synthetic stock market data by using a machine learning module to concoct order-book histories.

Separately, researchers at the Nanyang Technological University and the Chinese University of Hong Kong built a Gan to synthesize data for the S&P 500 and constructed an optimized portfolio using the information. The portfolios proved more resilient in real-world stress periods, they found.

And on the critical question of how to measure the realism of output data, Horvath and the quants at JP Morgan say they have made a breakthrough.

Mathematical signatures—a concept developed by Terry Lyons at Oxford University that originally was applied to read characters in written Chinese—can be used to encode the essence of a sequence of data, Horvath says. That allows models to measure precisely how similar one time series is to another. A

forthcoming paper on *Risk.net* will detail Horvath’s work with Lyons and the JP Morgan team.

Already, academics at University College London, the University of Kaiserslautern and the University of Edinburgh have developed a Gan that uses signatures in its discriminator. The method consistently beats “state of the art benchmarks” in forging realistic data, the authors state in a working paper released last year. And that includes Gans that learn based on a less sophisticated mechanism.

Several researchers are exploring models, or combinations of models, that reduce the noise in financial markets at the same time as learning to replicate it.

Horvath and her colleagues see promise in variational autoencoders, a type of model that compresses data into core elements, then rebuilds an alternative version following the blueprint the model has created. In its data generator, the team combined such a model with the discriminator element of a Gan.

Jhirad sees a future in allying autoencoders with Gans to help pick out salient features in markets and stop the Gan learning its way down blind alleys.

Amundi has been working with restricted Boltzmann machines, another type of generative model championed by quants at Standard Chartered in a study on synthetic data from 2019.

Boltzmann machines operate in a similar way to autoencoders, Roncalli explains, learning the probability distributions of interlinking elements of the data, from which synthetic data samples can be drawn.

The technicalities of faking data, then, remain to be pinned down. “Gans are really difficult to calibrate,” Roncalli says. “When you change the parameters, Gans are more sensitive. But we can’t say one method is better than another because this is a work in progress.” Quants are optimistic about succeeding, though.

Bigger datasets, a few human-imposed nudges in how the models learn, and neural networks with architecture built for the task at hand mean there’s a good chance of solving any teething issues, Jansen says. Gan-generated images of faces have come a long way from the blurry smudges first created in 2014. [Wt](#)



## 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](http://www.waterstechnology.com/newsletters) to update your newsletter preferences today

# Human Capital



## BNY Mellon Markets appoints head of financing and liquidity

BNY Mellon has hired Laide Majiyagbe in New York as head of financing and liquidity, assuming oversight of the firm's securities finance, liquidity services and collateral segregation businesses.

In this newly created position, Majiyagbe will oversee the strategic direction of BNY Mellon's financing, collateral and short-term liquidity offering and drive continued product integration between the businesses.

She joins after spending 14 years at Goldman Sachs in a variety of roles, most recently as global head of liquidity projections in the corporate treasury. She held several positions in liquidity and collateral management, and was involved in the design of a number of post-crisis collateral management and optimization tools used in managing firm and client collateral.

## Citadel Securities head of tech joins cryptocurrency exchange

Brett Harrison, former head of semi-systematic technology at Citadel Securities, has joined cryptocurrency derivatives exchange FTX.US as president. Harrison will work closely with the senior leadership team to establish



Laide Majiyagbe



Lorraine Waters

FTX.US as a major US-regulated cryptocurrency exchange.

Prior to his time at Citadel, where he oversaw a team of more than 100 engineers responsible for options, exchange-traded funds, American depositary receipts, and OTC technology, Harrison worked at Jane Street for three years, leading the firm's trading systems development group.

## State Street names head of Apac asset management clients

State Street has hired Neil Macdonald as head of its asset manager segment in Asia-Pacific. In this newly created role, Macdonald will lead the firm's engagement strategy in the region, including strategic direction and solutions structuring.

He joins from KPMG, where he was head of the Wealth and Asset Management Centre of Excellence, building a consulting business focused on wealth and asset managers in Hong Kong and mainland China. He has nearly two decades of experience in the asset management industry.

Based in Hong Kong, Macdonald will report globally to Donna Milrod, head of the global asset managers segment and global clients division, and regionally to Mostapha Tahiri, CEO for Asia-Pacific.

## AWS adds Perry Beaumont for finance data science

Amazon Web Services has hired former evaluated pricing specialist Perry Beaumont as a data scientist, focusing on areas including cloud computing, machine learning, and artificial intelligence. Beaumont will be responsible for identifying new insights into the uses and applications of AWS's analytic resources for finance.

He was most recently head of

data science at insurance provider Distinguished Programs, and has taught finance and data science at Columbia University.

## Broadway Technology names chief legal officer

Front-office solutions provider Broadway Technology has promoted Claudia Cantarella to the role of chief legal officer. She will also join the executive leadership team.

Cantarella began her relationship with Broadway as an outside advisor in 2008 while also sitting as a partner and co-chair of the New York intellectual property group of Salans LLP. She has been general counsel and corporate secretary at Broadway for eight years, establishing the company's in-house legal function and many of its internal policies, including information security, data protection, and compliance.

## Solidatus hires former HSBC executive as first CDO

Metadata management software company Solidatus has hired Lorraine Waters as its first chief data officer.

Waters will convene a new Solidatus customer advisory board and support the business in data privacy, environmental, social and governance, and financial crime risk management.

She previously held various roles at HSBC, including CDO for global compliance, financial crime and regulatory compliance.

## LiquidityBook hires Kost to lead sell-side business development

Software-as-a-service trading solutions provider LiquidityBook has hired Frank Kost in New York as head of sell-side business development for North America.

Kost will oversee business devel-



opment efforts for LBX Sell-Side, LiquidityBook's portfolio, order, and execution management system.

He was most recently director of business development at Axos Clearing. Before that, he was senior sales executive at SS&C. His previous roles include senior sales positions at SunGard and Aegis Capital.

#### **Capitolis appoints James Reilly as head of equity and TRS**

Capitolis has made James Reilly its head of equity and total return swaps (TRS) funding solutions. Reilly will lead the development and execution of a suite of equity financing products to help financial institutions address balance sheet needs and risk exposures.

Before joining Capitolis, Reilly was a senior managing director at State Street Global Markets and CEO of Currenex, a provider of FX technology owned by State Street. He previously served as global head of macro sales and trading at Cantor Fitzgerald.

#### **Gresham hires Cavell to lead North American bizdev**

Gresham Technologies, a London-based fintech firm specializing in data integrity solutions, payments and cash management, has hired Marc Cavell to its North American sales team as it looks to expand in the region.

Cavell has more than a decade of fintech experience spanning corporate banking, financial messaging, payments, and cash management. He previously held sales roles at Fundtech and Finastra, where he managed key accounts in the US market.

#### **Ex-BBH Stephen Bruel joins Coalition Greenwich**

Coalition Greenwich, a provider of data, analytics, and insights to the

## **BITVORE NAMES FORMER GOLDMAN EXEC AS CEO**

Alternative data firm Bitvore has appointed Elizabeth Pritchard as CEO to drive greater adoption among financial firms, and to expand the firm's offerings.

Pritchard has served as an advisor to the firm for the past 18 months. She takes over from Jeff Drake, who will continue to serve as president, focusing on partnerships and strategy.

Pritchard most recently ran her own alt data advisory firm, White Rock Data



**Elizabeth Pritchard**

Solutions. Before that, she was co-founder of go-to-market at Crux Informatics.

financial services industry, has hired Stephen Bruel as a senior analyst on its market structure and technology team.

Bruel will head the derivatives and FX practices, with an added focus on market infrastructure. He joins from Brown Brothers Harriman (BBH), where he was vice president and head of derivatives product management.

Before BBH, he was a research director at TowerGroup's securities and investment practice.

#### **EDI hires business development manager for Canada**

Market data vendor Exchange Data International (EDI) has hired former head of sales at GFT Group, Stephen Straker, for the role of business development manager, Canada.

Based in Toronto, Straker will enhance EDI's presence in the country's financial and FinTech communities.

He has experience in the Canadian market dating back to his early days as a trader on the floor of the Toronto Stock Exchange.

#### **SmartStream expands treasury business with new hire**

SmartStream Technologies has hired Peter Dehaan as new business director for cash and liquidity management.

Prior to joining SmartStream,

Dehaan spent seven years at Lloyds Bank managing treasury and liquidity services for various sectors under the global transaction banking umbrella. Prior to Lloyds, he spent almost 16 years at Citibank in various roles, including product management and correspondent banking.

Dehaan reports to Nadeem Shamim, head of cash and liquidity management at SmartStream.

#### **BCS Global Markets bolsters structured products team**

BCS Global Markets, the investment banking services division of Russia's largest independent broker, has appointed Konstantin Pavlov and Ekaterina Naumova as co-heads of its structured product team.

Pavlov will be responsible for structured product sales to Russian and international clients, while Naumova will focus on infrastructure development, including trading.

Pavlov joins BCS after spending three years at Rosbank, where he was head of structured products sales to institutional and retail clients. Naumova joined BCS as deputy head of corporate finance and later became the head of capital structures and finance. Before joining BCS, Naumova was the head of debt capital market at TransCreditBank. [Wt](#)



**Marc Cavell**



# The ESG Holy Grail doesn't exist... yet

As buy-side firms strive to stand out in a maturing ESG-driven market, Jo Wright says they will look for data in areas where coverage is still poor.



**N**ow that most buy-side firms offer investment products that have something to do with environmental, social, and governance factors—whether impact investment funds containing a mix of assets such as green bonds, or equities portfolios that exclude “sin” stocks—funds are looking to differentiate themselves. To do so, I predict they are going to want increasingly specialized ESG data.

For an article on ESG start-ups, in early May, I spoke to Paul Sinthunont, a senior analyst at Aite Group, who had just finished conducting a survey of about 50 asset managers from around the world, some very large and some boutique. Sinthunont told me that all these firms are looking beyond the ratings provided by the large data and index providers, such as MSCI, Morningstar, Bloomberg, and so on.

“The buy side doesn’t want to rely solely on ratings, even if they are a useful reference point. What they want is the underlying data so they can have their own view and build their own internal scores,” Sinthunont said. “They don’t want to rely on the same third-party providers as everyone else, because then they have no IP or thought process of their own. Many investment firms are trying to differentiate by promoting their ESG expertise.”

These firms would love a central platform where they could consume ESG data alongside their traditional financial data—that is the “Holy Grail,” Sinthunont says. But it doesn’t exist—at least, not yet. So, for now, they must be content with consuming data from at

least two or three vendors, if not more, all of whom plug the various gaps that remain in ESG data.

As *WatersTechnology* has reported, many of these gaps are the result of inconsistent corporate reporting. In developed markets, this reporting has improved a lot: recent research by S&P Global found that in 2019, 90% of the biggest companies in the US published sustainability reports; in 2011, it was just 20%. This is partly a response to investor demand, and partly regulatory requirement. The EU’s Sustainable Finance Disclosure Regulation, for example, requires investment firms to source lots of granular data on companies to calculate the impacts of their investments.

Even so, corporate reporting remains voluntary in many cases, and opaque. In asset classes like real estate or private markets, there are very few ESG providers and no data standards. In emerging markets, ESG is in its early stages. For large buy-side firms, the data to measure risk and assess performance in multi-asset portfolios is unavailable from the large data providers that dominate the ESG data market.

## Raw ideal

Buy-side firms attempt to fill these gaps themselves. *WatersTechnology* has spoken to asset managers about their ESG integration processes, and while they are very different, they seem to follow the same basic philosophy: the best approach is raw data from corporates, combined with data from vendors, then compared against internally developed metrics or scorecards, and consumed

on internally developed platforms.

Most recently, we wrote about how some funds are eyeing raw data sourced from non-profits like environmental activist groups to inform their investment decisions, rather than relying on black-box ratings from data vendors. The fund managers want to draw their own conclusions, and believe that to do that, they need the raw data from corporates that underpins their ratings, bolstered with data from the non-profits.

As Mike Chen, director of equity and head of sustainable investments at PanAgora Asset Management, told us, “There’s no right or wrong; it’s like asking someone what their favorite color is. I think therein lies the problem of using commercial ratings—you’re accepting somebody else’s decision for you.”

Chen says these types of data are more useful than pre-packaged information from an ESG ratings agency when compiling an ESG-based portfolio. “If an NGO puts out a report that a certain company has bad employment practices, or a company has been polluting, we read that data and then decide how important or relevant those reports from the NGO are to the way we view ESG,” he says.

We often hear how alternative data can supplement more traditional sources, whether for fundamental or quantitative investment strategies. And real estate, private markets, and emerging markets are some other gaps that data vendors—whether large ones or start-ups looking to exploit a niche—will be looking to fill as investment firms race to retain their competitive edge. [Wt](#)

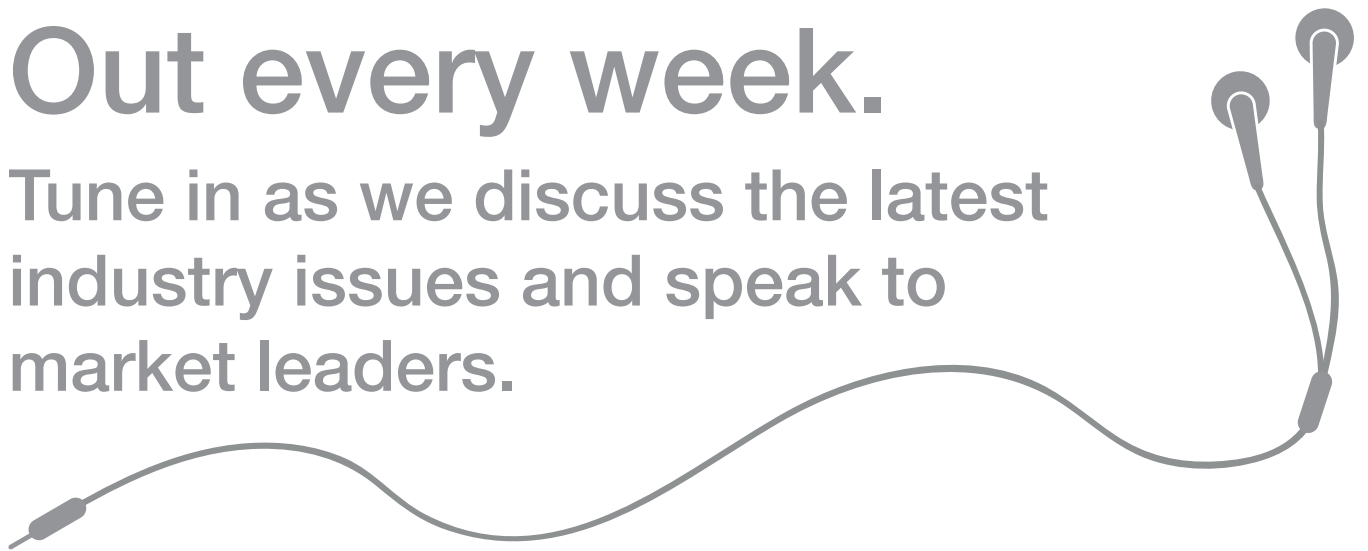
**waters  
tech**



# Waters Wavelength Podcast

Out every week.

Tune in as we discuss the latest industry issues and speak to market leaders.



Subscribe at:  
**[soundcloud.com/waterstechnology](https://soundcloud.com/waterstechnology)**  
or search the iTunes store for  
**“WatersTechnology”**





# 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](mailto:info@smartstream-stp.com)**  
**[smartstream-stp.com](http://smartstream-stp.com)**