Inside Data Management
Inside Market Data Inside Reference Data

The Big Questions: Ethics and Alternative Data

FIRST SHOTS FIRED IN MARKET DATA FEES BATTLE

You Down with CTP? Yeah, You Will Be

SEMANTICS' SEMINAL CONVERGENCE

Data Ontology Breakthroughs and Creative Collaborations Prove Pivotal

Inside Market Data Inside Reference Data Buy-Side Technology Sell-Side Technology
waterstechnology

Follow Topics

Build your own email alerts

Follow our coverage on individual organisations like ESMA or ISDA. Or pick specific topics such as Mifid II, blockchain, or artificial intelligence. Plus many more. Select your topics and you'll receive everything you are interested in, all in one email.

Visit waterstechnology.com/follow

Editor Jamie Hyman
Tel: +44 (0) 207 316 9270
jamie.hyman@infopro-digital.com
Asia Editor Wei-Shen Wong
Tel: +852 3411 4758
wei-shen.wong@infopro-digital.com
European Reporter Amelia Axelsen
Tel: +44 (0) 207 316 9074
amelia.axelsen@infopro-digital.com

Publisher Katie Palisoul
Tel: +44 (0)20 7316 9782
katie.palisoul@infopro-digital.com
Managing Editor Max Bowie
Tel: +1 646 490 3966
max.bowie@infopro-digital.com
Business Development Executive Arnaud Morell-Coll
Tel: +1 646 736 1887
arnaud.morell-coll@infopro-digital.com

Group Publishing Director Lee Hartt
Head of Editorial Operations Elina Patler
Production Editor Brett Gamston
Commercial Editorial Manager Stuart Willes

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

Infopro Digital Head Office
Haymarket House
28–29 Haymarket
London SW1Y 4RX
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

Subscription Sales
Tsetso Bikov
tel: +44 (0)20 7316 9326
tsetso.bikov@infopro-digital.com

Infopro Digital Customer Services
E-mail: customerservices@infopro-digital.com
Tel (UK): +44 (0)1858 438800
Tel (US): +1 212 776 8075
Tel (Asia): +852 3411 4828

Inside Data Management (ISSN 2514-0574) is published monthly (12 times a year) by Infopro Digital Risk Limited. Printed in the UK by DG3.

Shaking Off the Semantics Stereotype

When you write about semantics, a funny thing happens: Sources get self-deprecating. For my research into the industry's march toward game-changing semantics ontologies, I spoke with some of financial services' most insightful minds, certainly—people who are excited by the topic, and nearly every one of them made a joke about how semantics is terribly dull or remarked how no one ever wants to talk about standards or the considerable, complex work that goes into establishing workable ontologies.

They're wrong, though. At this very moment, the semantics world is forming creative partnerships and producing ground-breaking technologies and although so far there is only a glimpse of what could be gained by establishing and adopting ontologies across the industry, that glimpse is exciting. For this issue, I took a first look at how, after years of work, the necessary convergence of standards, ontologies and their respective custodians is finally becoming a reality. In fact, I found the topic so robust that I plan to expand it into a series, so look for features in upcoming issues of *Inside Data Management* that focus on other dynamic breakthroughs in the world of semantic ontologies.

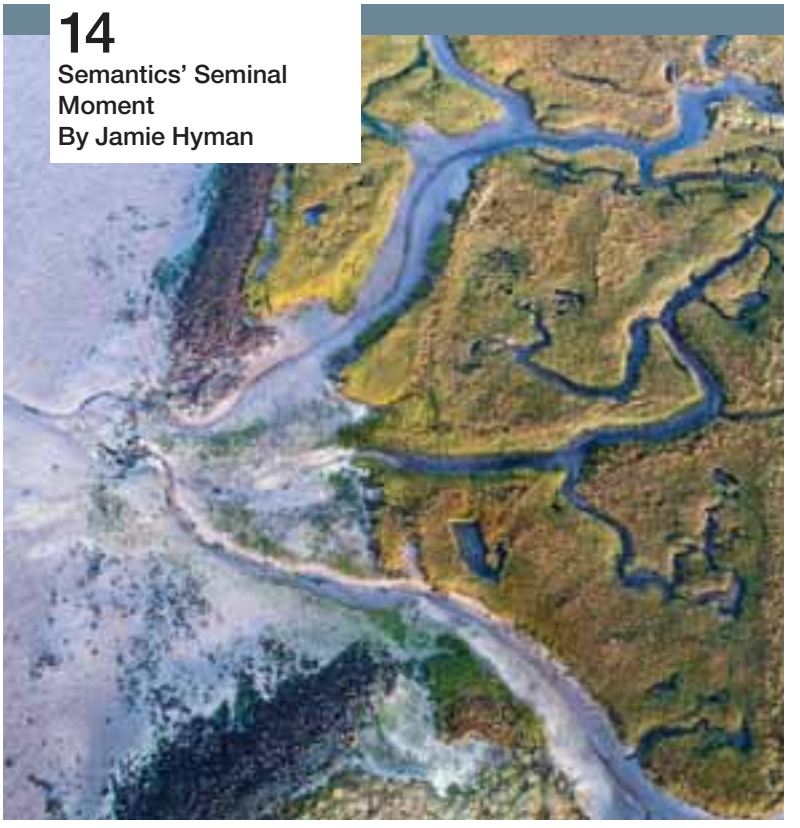
Our Hong Kong editor, Wei-Shen Wong, also takes a closer look at how data is handled. We often hear data managers struggling to stay ahead of regulatory requirements mutter that although they mandate reporting, regulators aren't equipped to handle all that data, anyway. Inspired by a European Securities and Markets Authority keynote speaker at the European Financial Information Summit in September who said the regulators are handling the data just fine, thank you very much, Wei-Shen spoke to their Asia Pacific-based and Japanese counterparts to learn more about what the region's regulators do with the data they receive, and how they plan to improve their own processes.

Amid various discussions about how data is organized and managed, it is always worth taking a moment to mull whether the data should be eligible to deliver alpha in the first place. Amelia Axelsen's feature on the ethics of alternative data investigates why and when firms should consider data ethics (hint: early) and reveals what are the big questions to ask to head off an ethical day of reckoning. Another area due for a day of reckoning is the rising cost of market data. Max Bowie provides a preview of industry user group and Nasdaq positions ahead of a two-day Securities Exchange Commission event slated for late October. ■

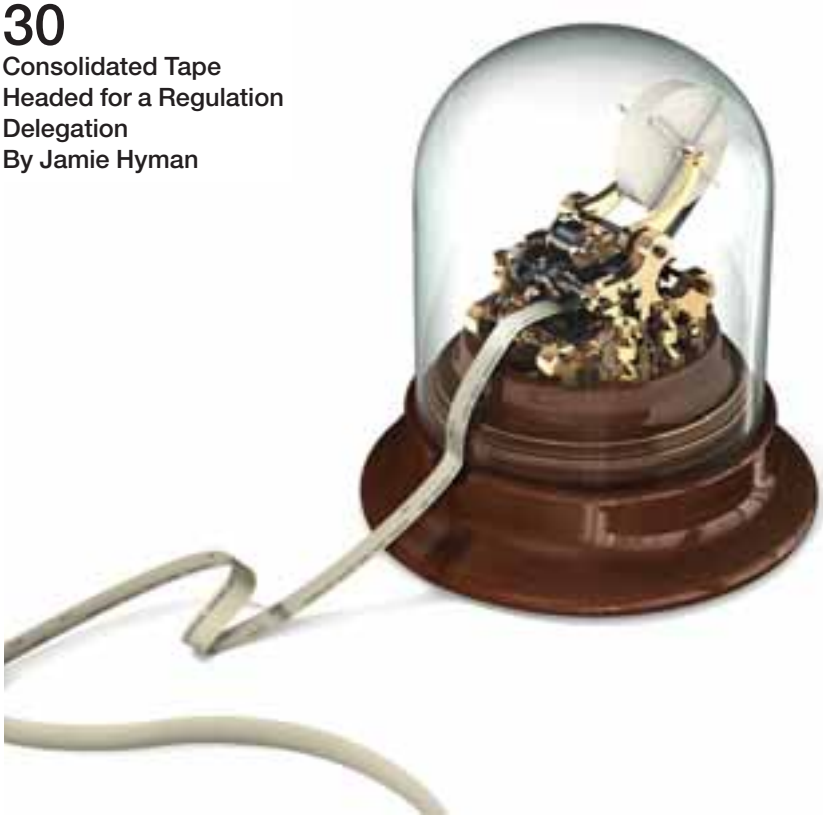
Jamie Hyman
Editor

Inside Data Management
Inside Market Data Inside Reference Data

Contents



October 2018



waterstechnology.com
For more information and readers' feedback please join the discussion

HSBC Expands Content of Research App on Symphony

HSBC's Global Research division has added video content to its Research Application after recently rolling it out to clients via Symphony Communication Services' electronic chat platform, and is planning future upgrades to take advantage of artificial intelligence to broaden the app's appeal.

HSBC's app is one of the most popular on Symphony

The bank made the research app available on Symphony for internal use around a year ago, and rolled it out externally to clients on Symphony a couple of months ago, according to officials. Since then, it has already become one of the most popular apps on Symphony, with roughly 10,000 individuals who have access, of whom "thousands" of internal HSBC staff and external clients access it daily, said Alex Andronov, head of business development at HSBC Global Research, speaking at the Symphony Innovate 2019 conference in New York.

HSBC has added the ability for research analysts to record video commentary at their desk, which then—after an internal approval process—is integrated in the research app.

Andronov said the bank first rolled out the app on Symphony—which is the principal internal chat system used by the Global Banking and Markets division, though not the whole of HSBC—for internal purposes, but has seen increased readership of its research as a result of making it more widely available.

"Initially, we didn't think it would be client-facing… [but] in each channel we add, a significant number of people read the research who haven't accessed it before," he said. "Our plan is to go to where the customers are…. People want research to come to them. They want it available in chat because they are using chat… and we want to make it as frictionless as possible [to access our research]."

Future versions of the app software will include artificial intelligence that understands the types of research and topics that clients are reading, and is able to recommend other research items that they might find interesting. However, Andronov said the firm is unlikely to use that information in direct client communications, but will make it available to its salespeople.

"We're not just saying to end-user clients, 'Here are some things you might be interested in.' We do have some AI things like that, but we're going to make that available to salespeople first, for them to use in their communications with clients," he said.

Selerity Touts Improved Search, Less 'Chatter' for Crypto Data

New York-based contextual search and analytics provider Selerity is rolling out an update to its Selerity Context engine that searches public and private information sources for information relating to the top 20 cryptocurrencies by market cap.

The cryptocurrency support can search data sources such as news, research and market commentary—plus social media via a Twitter integration feature—as well as a client firm's internal datasets and digital communications, such as chat programs and emails.

Selerity CEO Ryan Terpstra says the new feature moves away from a purely keyword-based search to one that is concept-based, reflecting both the evolution of cryptocurrencies from a purely retail phenomenon to a hot topic among institutional investors, adding that while investors want to monitor the drivers of performance for cryptocurrencies against one another, they also want to monitor themes such as regulatory developments, and updates to the underlying technologies relevant to crypto markets, such as Blockchain.

"This is about being able to take large volumes of content relevant to cryptocurrencies and present it to clients. If you do a search for Ripple, for example, you get a lot of low-value content. We filter out a lot of the chatter so you can hone in on just the important information," Terpstra says.

DSB Announces New Services, Fee Discounts for 2019


Sassan Danesh, Etrading Software

The Derivatives Service Bureau (DSB), the utility created by the Association of National Numbering Agencies (ANNA) to generate International Securities Identification Numbers (ISINs) for OTC derivatives, has announced user fee rebates in 2019, plus enhancements to the OTC ISIN functionality to enable real-time updates to DSB reference datasets planned for later next year.

Sassan Danesh, managing partner of DSB's technology partner Etrading Software, notes that DSB's platform—which runs mainly in the cloud—is fully automated and managed in near real-time, and generates more than 99 percent of ISINs in less than one second. When creating the solution, Etrading Software had the "luxury of a clean slate," unburdened by outdated legacy systems.

DSB is a non-profit utility running on a cost-recovery model. Firms are required to use an ISIN if an entity trades OTC derivatives with a European counterparty or with a European trading venue.

"When we [DSB] went live, we were conscious that there were legal requirements for European entities to utilize the ISIN, so we wanted to keep costs as low as possible," Danesh says.

DSB has further upgrades planned for 2019.

IHS Markit Unveils Initial Margin Calculator to Aid ISDA Margin Rule Compliance

IHS Markit has released an Initial Margin Calculation Service to help investment firms calculate and post accurate and compliant amounts when trading over-the-counter derivatives.

The service currently supports credit default swaps, interest-rate swaps and vanilla equity products, and will soon support currency forwards and other products, as part of an "aggressive" ramp-up to support most instrument types by the end of this year, says Laura Misher, managing director of derivatives data and valuation services at IHS Markit.

The vendor rolled out the service in June, based on its existing valuation services for non-cleared interest-rate, currency, credit, equities, and commodities derivatives—which the vendor has offered since 2005, and which are used by more than 400 direct clients and "thousands" of indirect clients (such as those who receive valuations via fund administrator clients) to support net asset value reporting and margin calculation—and expects to go live with its first client by year-end.

Misher says the service was driven by client demand—primarily from fund administrators that offer the service on behalf of clients and who already use Markit for valuations but are looking to offload that requirement ahead of new rules governing the posting of initial margin from the International Swaps and Derivatives Association that will take full effect in 2020.

Laura Misher, IHS Markit

"Implementation of the ISDA SIMM (Standard Initial Margin Model) was a milestone for us," she says. "We had to license it, be approved, and make sure that we keep up with any changes to the model."

Markit offers Initial Margin as a managed service, running in Amazon Web Services' cloud. Clients send their positions data to the vendor, and it sends them back initial margin calculations as a feed or file. Existing clients of Markit's valuations services don't even need to re-send their data.

But that doesn't mean creating the service was without challenges: "We ran into hurdles around scale of processing because so many calculations need to be run on a position level. We were running into instances where it was taking a long time to complete," Misher says. "So we implemented an advanced algorithmic approach [Adjoint Algorithmic Differentiation (AAD), which is already widely used to calculate banks' valuation adjustments risk] that allowed us to speed up the calculations in a scalable fashion."

Arabesque Adds Eagle to List of ESG Data Users

Arabesque, a quantitative fund specializing in sustainability, has partnered with Eagle Investment Systems to offer access to its environmental, social and governance (ESG) data service, S-Ray. The S-Ray sustainability data tool, which Eagle clients can now incorporate into their platforms, applies ESG metrics using 50,000 sources in 15 different languages to monitor thousands of the world's largest corporations. Arabesque excludes companies from its portfolios if the metrics give them low sustainability rankings. The shares of some excluded companies, such as Volkswagen and Facebook, suffered in the wake of scandals, sparking the interest of investors and driving demand for ESG data from asset managers, says Andreas Feiner, CEO of Arabesque S-Ray. He adds that the company's core mission is "getting sustainability into the mainstream" investment process. "Arabesque has an interesting approach because they are both an asset manager and a data provider—in other words, they eat their own cooking," says Joel Kornblum, head of strategic alliances at Eagle.

TRG Screen Expands ResearchMonitor Data Tracking Tool

Data inventory monitoring software vendor TRG Screen has expanded the capabilities of its ResearchMonitor web data tracking tool to include locally installed data and software applications, allowing clients to monitor web- and application-delivered services via a single solution.

ResearchMonitor—which TRG acquired with the purchase of Priory Solutions in 2016—was previously used by legal and financial firms to track usage of web-based information sources. The update allows firms to expand its use to other data sources.

"Market data managers have been so good at providing transparency into data subscription costs that they are now being asked to manage other areas—and the next-largest costs are software subscriptions for desktop applications," says Richard Mundell, CTO at TRG Screen. "There are plenty of items out there—both market data and non-market data—that are only available as desktop applications."

Richard Mundell, TRG Screen

The vendor already offers the DART desktop usage monitoring tool—which it sells under an exclusive license from owner Vela—though Mundell says there are key differences between the two.

"The distinction between ResearchMonitor and DART is that DART is very specific and very deep—it tells you the content being looked at, whereas ResearchMonitor tells you more about how long someone has been looking at or interacting with a product," he says.

Arachnys Raises $10M to Fund KYC Risk Tools Expansion

London-based Arachnys, a provider of customer risk data to support Know Your Customer and Anti-Money Laundering requirements, has raised $10 million in Series A funding to accelerate product development, go-to-market efforts, and strategic partnerships.

Regulatory pressures are fuelling demand for KYC/AML data

Arachnys president Ed Sander says the vendor will accelerate the product roadmap of its Customer Risk Intelligence platform, which comprises the Spotlight, Navigator and Sentry tools for building and enriching existing institutions' client profiles, incorporating third-party information, and automating profile risk updates and entity relationships, respectively. One specific initiative speeds up the development of machine-learning capabilities to help analysts reuse investigations "to work better, faster, and smarter," he says.

The cash will also support Arachnys' pursuit of partnerships with other financial crime solutions providers and consultancies—in addition to existing relationships with Fenergo and Quantexa—and broader relationships with analysts, and enable it to hire additional salespeople, and will help fund the recent formation of a new delivery and implementation organization created to meet demand, which constituted a "significant" increase in staff in London and New York.

"We've had significant expansion in our pipeline, and we are expanding to meet this demand," Sander says, adding that the demand is a result of greater regulatory pressures.

"There are many new regulations that have come out since 2015 that focus on new requirements for banks to: (1) obtain information to verify the "nature of a business relationship" between parties involved in transactions; (2) define "beneficial ownership structures and relationships for companies and individuals…; and (3) obtain at least two separate public sources to verify information obtained about a company—e.g., a corporate registry to confirm basic demographic details about a bank's place of incorporation, business registration, etc.," Sander says.

For example, he cites a deal to help a tier-one European bank to streamline its client onboarding process by rolling out Arachnys to more than 2,000 risk investigators.

"The bank is using Arachnys as a single, global KYC research engine to simplify client outreach and avoid asking repeat questions during the onboarding process," he says. "By utilizing a cloud platform, investigators around the world are able to work in parallel (and in real time) with the relationship managers dealing directly with customers. The firm envisions this will trim at least two weeks off its standard onboarding process every single time for every new customer."

The funding was led by QED Investors, a venture capital firm focused on high-growth information companies. As a result of the investment, Amias Gerety, a partner at QED, has joined the Arachnys board. Sander says that as acting assistant secretary for financial institutions at the US Department of the Treasury under both Obama administrations, Gerety has direct knowledge of the regulations in question.

Bloomberg Unveils Online Portal for Normalized API Data

Gerard Francis, Bloomberg

Data giant Bloomberg has launched an online data portal of normalized pricing and reference data, dubbed Enterprise Access Point, which is designed to make it easier for data scientists and artificial intelligence-based applications to consume consistent and standardized historical data via an API rather than traditional bulk file transfers.

Clients of Bloomberg's Data License service can access the data—which comprises a super-set of the vendor's data, minus its real-time datasets, including pricing, reference, regulatory and historical data—at eap.bloomberg.com.

Gerard Francis, global head of enterprise data at Bloomberg, says the vendor created EAP based on feedback from financial firms that they wanted easier ways to access the data in a normalized format.

"[Clients] want data that's really easy to consume. And the more points of friction we can eliminate, the better," Francis says.

TPI Partners with Yonhap to Distribute OTC Data

Real-time price information provider Tullett Prebon Information (TPI) has expanded its partnership with South Korean real-time news and financial data provider Yonhap Infomax.

TPI now distributes and licenses its over-the-counter (OTC) data via Yonhap's data feed. Rhys Spencer, regional sales manager for Asia at TPI, says demand for OTC data in South Korea is on the rise, driven by securities firms issuing more innovative products.

"In addition, new sourcing requirements at financial institutions, including reviews of existing providers and demand for international data among domestic banks, are factors increasing the use cases for OTC data," he says.

Spencer says TPI and Yonhap have been partners for more than a decade, previously on Yonhap's desktop data provision.

This latest expansion provides South Korean domestic banks and securities firms with a new option for accessing TPI's OTC data for swaps, fixed income and foreign exchange instruments.

JCV Vendor Audit Module Tackles Data Under-Reporting

Colchester, Conn.-based market data audit specialist JCV Investment Systems has released a vendor audit module as part of its Genie data audit platform, which officials say will allow exchanges to significantly improve the speed and accuracy of audits conducted on vendors that redistribute their data.

The new module automates the process of collecting and analyzing usage and entitlements information relating to exchange data sold by vendors, and providing reports that identify instances of under- or over-reporting, where a client firm may be using more data accesses than the vendor reports to the exchange.

"Most exchanges aren't interested in chasing after their vendors because it's hard to do and takes a long time," says Joe Veneziano, global markets director at JCV, which performs data audits on behalf of exchanges. "So this will enable them to do it quicker, cheaper, and more accurately. Genie means they have a tool that reduces the amount of work associated with this by 90 percent. It will help them find vendors who are doing a bad job of reporting—and then the exchanges might want to look at the clients of those vendors."

"Most exchanges aren't interested in chasing after their vendors because it's hard to do and takes a long time. Genie means they have a tool that reduces the amount of work associated with this by 90 percent." Joe Veneziano, JCV Investment Systems

JCV previously performed this process manually for exchange clients. Now, using the new module, Genie captures vendors' audit trail information from spreadsheets, CSV or .txt files, converts them into its own format, and identifies any discrepancies between that and the usage they report to an exchange, then delivers a detailed report showing individual end-users and the duration of their usage of each product.

"The exchange knows who it is selling data to, and the vendor reports who it is selling to, with user contracts. Then we go to the vendor and get their audit trail… of who at each firm had access each month. We compare the counts from the exchange at the subscriber level with the report from the vendor that has the name of the client company and individual users for any given month," Veneziano says. "So we can not only show any under-reporting, but can also show the vendor where they went wrong, so they can fix it. Feedback from vendors has been that this is very helpful—they don't like to be wrong."

Veneziano says that while JCV does uncover some instances of what he believes to be intentional under-reporting of usage, most errors are a result of poor processes at vendors—for example, not including subscribers who are added in phases after an initial rollout, or as a result of a misunderstanding of issues such as licensing rules governing professional versus non-professional usage, netting, or unit of count.

Survey: Data Residency a Key Concern for Cloud Adoption

Brennan Carley, Thomson Reuters

When it comes to migrating data to a public cloud, firms are concerned about where to put the data, who can see it, and who has control.

A survey conducted by Thomson Reuters Financial and Risk—the company's service branch—reveals that 24 percent of respondents cite data residency, 19 percent cite data privacy, and 18 percent cite concerns about losing control over data when switching to public cloud.

The onus is largely on firms to ensure that data privacy and residency measures are negotiated and outlined in contractual agreements with cloud providers, says Brennan Carley, global head of enterprise proposition and product for the Financial and Risk business at Thomson Reuters.

"Firms… understand that just because the tools are there and the cloud providers have done a good job at securing their infrastructure, it's still up to the banks [and] asset managers to ensure that they're honoring the different regulations around data residency," Carley says.

FESE Requests Clarification on LEI Application

As Legal Entity Identifier use becomes more widespread, nagging application inconsistencies remain.

In feedback to the Financial Stability Board, the Federation of European Securities Exchanges (FESE) has highlighted problems—specifically, inconsistencies in how LEIs are applied to some entities.

"We've asked regulators for more clarity as there doesn't appear to be a consistent approach to the application of LEIs to certain entities such as investment funds," says Rainer Riess, director-general of FESE. The federation's feedback cites as an example that "the application of LEIs to investment funds varies as some funds have LEIs at umbrella level, while others have LEIs at the sub-fund level."

Stephan Wolf, CEO of the Global Legal Entity Identifier Foundation (GLEIF), says GLEIF does not comment on the application of LEIs determined by authorities in individual jurisdictions.

"With regard to the FESE observation, it should be noted that the LEI Regulatory Oversight Committee (LEI ROC) is working on guidelines for the handling of funds," Wolf says. LEI ROC did not respond to a request for comment.

Nasdaq, User Groups Fire Opening Salvoes in Battle Over SIP Data Fees

As the SEC prepares to host a two-day event to tackle market data access and fee issues, industry user groups and Nasdaq have laid out their positions on SIP reform. Max Bowie reports.

In late August, two industry user groups—the Managed Funds Association and the Alternative Investment Management Association—penned a joint letter to US markets regulator the Securities and Exchange Commission, requesting reviews of the fees charged by US exchanges for market data.

Though primarily concerned with the cost of proprietary exchange data, the letter asked the SEC to apply the same principles to governance of the SIP (Securities Information Processor) feeds—the consolidated feeds of US quote and trade data collected and distributed on a utility-like basis by Nasdaq (the UTP program) and NYSE (the Consolidated Tape Association), respectively.

For example, the associations want the SEC to review market data fees to see whether the prices charged correspond with the cost of operation, and to be more rigorous when approving proposed fee changes associated with new products.

In the letter, they describe how regulatory and technological developments have contributed to greater competition, market access and liquidity, more market depth, tighter spreads, and lower costs, but note that while the cost of technology has become cheaper, market data costs continue to rise as exchanges increase fees and unbundle products into separately billable services. At the same time, the usefulness of the SIP feeds has largely been overtaken by direct exchange feeds, which the associations say firms feel a fiduciary obligation to buy in order to obtain the levels of data they need to support their strategies, urging the SEC to reconsider whether firms "can truly meet their best execution obligations relying solely on SIP data feeds."

Nasdaq wants dialog between SEC, exchanges and investors

So in addition to urging a cost-based review of SIP data, the associations also called for an investigation into the governance structure of the SIP feeds to mitigate any conflicts of interest with the exchanges that contribute to them.

Nasdaq, for its part, says it wants to urge the SEC to clarify its position on the vendor display rule, which officials say has created confusion among broker-dealers over when they may use SIP data instead of faster, more comprehensive and more expensive direct exchange feeds.

Shortly after the MFA and AIMA published their letter—though officials say not in direct response, but rather in anticipation of a roundtable discussion on market data and market access being organized by the SEC on Oct. 25 and 26—Nasdaq released its own set of proposals for modernizing the SIP. In addition to the vendor display rule, Nasdaq's proposals include changing the SIP revenue formula to incentivize "lit" quotes, and expanding the authority of the SIP Advisory Committee.

"We think the investing public in general benefits from lit quotes. There has been a growth in 'dark' trading, and we would like a change in the exchange revenue allocation to recognize that lit quotes enable pre-trade price discovery… and improve the level of pre-trade information out there," says Oliver Albers, SVP and head of strategic partnerships for Nasdaq's Global Information Services business.

Continuing its stated aim of supporting the investing public, Albers says its plans to expand the authority of the SIP Advisory Committee are not about giving it free rein to increase fees, but rather about opening up involvement to the investing public.

"We believe there are benefits in having more industry representation at the committee level," Albers says.

Here, at least, the parties may be on the same page: As the MFA and AIMA note in their letter, "as non-voting members, the advisory committees have very little authority to influence issues of major impact to retail and institutional investors, broker-dealers and other market participants, such as SIP fee increases and technology upgrades. As a result, the advisory committees to the SIPs are not able to mitigate the conflicts of interest that exist between exchanges charged with overseeing the SIP data feeds, and that in turn sell proprietary market data products and connectivity offerings… that potentially compete with the SIP data feeds."

But Nasdaq is wary of further structural change to the SIPs' operation. "A lot of people are advocating for a distributed SIP… which would change our market structure. Nasdaq's position is that we think the structure of the SIP today works pretty well," Albers says. "Multiple or distributed SIPs could be confusing and benefit some stakeholders at the expense of the investing public… and may mean a significant increase in costs—who would bear those costs?" ■

French Regulator Says Mifid II Data Quality 'Difficult'

AMF chairman Robert Ophèle says data quality and completeness pose problems for regulatory evaluations of Mifid II transparency requirements, and that regulators will review frameworks following Brexit. Amelia Axelson reports.

The chairman of the Autorité des Marchés Financiers (the French Financial Markets Authority, or AMF) says there are lingering problems with the implementation of Mifid II and has warned about data shortfalls ahead of possible deregulation resulting from the UK's exit from the European Union.

Robert Ophèle
AMF

At the Annual European Compliance and Legal Conference (AFME) in London, AMF chairman Robert Ophèle said it is up to firms to examine key regulatory frameworks following Brexit and to ensure they are reporting higher quality data.

The French regulator highlighted how market structures have changed under the sweeping European regulation with an influx of information being generated by market participants, but told delegates the data still poses problems for effective evaluation of risk metrics by European regulators.

"The first nine months of [Mifid II] implementation have highlighted just how rapidly markets can shift, how essential level 3 guidance is, and how difficult a task it can be to ensure data quality," he stated.

Ophèle said while innovation is championed in the market, some structural changes have arisen not for innovative purposes, but as a way to circumvent Mifid II regulatory frameworks. One example he provided was the shift of European commodity derivatives contracts from European venues to US venues, or the switch from regulated venues to over-the-counter.

To improve technical support and reach a "level playing field" in the European Union, he called on the convergence between National Competent Authorities (NCAs) to provide fundamental, uniform level 3 guidance to participants. While the European Securities and Markets Authority's Q&A's and opinions provide governance, many NCAs don't follow the regulator's guidance or implement the proper tools under EU law, according to Ophèle.

Mifid II has been demanding on both firms' and regulators' technology frameworks, and so technical guidance has been necessary, but data—"the core" of Mifid II's transparency regime—has been lacking, he said.

"Mifid II is also a sizeable IT challenge, indeed perhaps an unprecedented data project both for market players and market regulators; this data is at the very core of Mifid II implementation since all transparency requirements hinge largely upon it," Ophèle said, adding that the Financial Instrument Reference Data System run by Esma manages no less than 12 million ISINs, that 1.3 million LEIs have been issued worldwide and around 500 million transaction reports are exchanged between NCAs on average every month through the Transaction Reporting Exchange Mechanism. "But while the amount of data requested is enormous, the quality of this data still leaves much to be desired," he said.

Although EU financial instruments' reference data quality is monitored, Ophèle noted that regulators are receiving inconsistent data and incomplete declarations. One example was the implementation of the double volume cap (DVC) in the first quarter of 2018. In a report on September 27, Esma, in conjunction with NCAs, deemed completeness "unsatisfactory" in regards to the data underpinning the monthly DVC and quarterly bond liquidity assessment publications. The report provided details on two new indicators—the completeness ratio and completeness shortfall—for closer inspection of trading venues' data, and noted steps to improve data quality across the European market.

"Looking back over the past nine months, I believe there is a growing awareness that a number of legislative fixes will need to be considered in the short to medium term to correct certain deficiencies," said Ophèle. "It will render all the more legitimate a number of recalibrations or redrafts to be included in this exercise, some of which may actually stem from Brexit's consequences."

He expects that re-examining some of the regulatory frameworks under Mifid II may be necessary following the final deliberations of the UK's exit from the European Union.

"The industry must be fully aware that there is a strong commitment from Esma and all EU27 NCAs to have an MMoU with the UK FCA in place. It is upon this assumption that market players should work when putting their contingency plans in place," said Ophèle, adding that reviews could focus on narrowing the scope of trading obligations and evaluating third-party issuers that could be simplified post-Brexit. ■

Japan Exchange Group Mulls Pooling Content into Data Lakes

Wei-Shen Wong talks with the CIO of the Japan Exchange Group about why he is leaning toward data lakes when considering JPX's data management approach.

As exchanges expand the amount of data they collect and distribute—which increasingly involves new data types that do not conform to "traditional" market data structures—they are revisiting the potential of data lakes to deliver the breadth and depth of capabilities to store and surface both structured and unstructured data at any scale in the most efficient manner.

One such potential exchange user, the Japan Exchange Group (JPX), is looking at the potential of data lakes to support its strategy of finding new ways to add value to existing proprietary data.

Data lakes are centralized repositories that store data in an ungoverned and unformatted way, allowing firms to deploy tools to run different types of analytics and machine-learning applications without having to move the data. Used and implemented correctly, they could help organizations yield significant improvements in their ability to store and manage certain datasets, particularly when it comes to experimenting and building out potential new products.

It's this aspect that particularly appeals to JPX CIO Ryusuke Yokoyama, who says that—while not yet set in stone—the exchange is currently thinking about using a data lake to help achieve this.

"At the moment we only have a data warehouse. We don't have a data lake or data pool where we can inject all the data. We are thinking about it but we haven't realized it," Yokoyama says.

Ryusuke Yokoyama
Japan Exchange Group

A data warehouse stores structured data differently from a data lake, which organizes data while accounting for business processes and determines how data sources are analyzed. When a purpose for the data has been identified, only then will it be loaded into the warehouse.

According to advisory firm Deloitte, data lakes should be divided into three zones: data loading, defining user access and security, and creating a more user-friendly environment. The first zone consists of the raw and untransformed data direct from the source. Zone two is the data sandbox, where the data can be lightly processed, cleansed and combined for exploration and analysis. The third zone consists of the refined data that is ready to be used by the data warehouse or manipulated for analysis.

JPX is currently replacing its internal systems, including the system that accumulates data, Yokoyama says. "We are constructing a system considering the convenience of data usage and the smoothness of cooperation with other systems," he says.

Like other organizations, JPX is trying to gain more value out of the data it already has in order to better cater to its customers and market participants. Adding to this, artificial intelligence (AI) and machine-learning technologies are also becoming more sophisticated. "So we believe that data utilization will expand more in the future. JPX holds various data, so I think that what kind of value we should add to raw data is the issue for us," he says.

But while the exchange sees the opportunity presented by data and is considering ways it can expand its data business, exactly how it will do so is still uncertain at this point.

"Everybody is scrambling to the data business. I think it has huge potential in JPX's future growth and expansion. What we're doing right now is providing the plain and simple raw data, like stock prices and volumes. What we have to think about going forward is how we can add value to that data and sell it to customers, and how we can work out their needs and appetite for data," Yokoyama says.

That may even mean that JPX will need to direct some efforts towards development of AI technology "to filter and pick out necessary data," he adds.

Though deciding how JPX can add value to its data and repackage its data as new products is a business decision, rather than one that will be determined by the data and IT teams, the IT team still has its work cut out. "On the IT side, we have to think about how we can handle the massive data. How can we build the data handling platform that can respond to the business needs?" Yokoyama says. ■

Virgin Atlantic's Data Strategy: The Offense Funds the Defense

Jamie Hyman reports on how Virgin Atlantic established its data program, as outlined by the airline's data chief at the European Financial Information Summit.

One short year ago, Virgin Atlantic's data strategy was just getting off the ground.

"In 2017, we started off with no data warehouse," Tim Lum, the airline's head of data and insight, told delegates at the European Financial Information Summit. "Data was in complete silos. We didn't even have an enterprise data warehouse. In similar vein to regulated industries, we had them in specific silos in order to cater for specific regulations. We had no governance, no process, no analytics. We only had basic reporting across our business."

Tim Lum
Virgin Atlantic

Virgin Atlantic has since upgraded. "Where we have put our money where our mouth is today is to really invest in data and digital platforms, and innovation, as one of the only ways to prevent ourselves from being further disrupted and, therefore, protect ourselves from what we're actually seeing in the market today," Lum said—namely, disruptive low-cost carriers that have forced legacy airlines to unbundle services and be more nimble.

He told delegates the first steps in the journey from fragmentation to innovation consisted of technology, then people, and process.

"We had to start with technology. We didn't have any capability for our individuals to grow, to learn about data management, to start implementing models, because we just didn't have the infrastructure… We were using our ERP [enterprise resource planning] systems as our central data warehouse and that wasn't working," Lum said, adding that Virgin Atlantic is building a cloud-first data platform, and that "We are not building any more datasets in on-premise data warehouses."

Once the airline moved to the cloud—it uses Microsoft Azure—it turned its attention to its people and processes; specifically, thinking about how to implement artificial intelligence (AI).

"The basic premise of all our AI algorithms and models are to land them with human supervision, and that includes anyone who hasn't really touched data before to all of our data scientists today," Lum said. He told delegates that Virgin Atlantic's first model was an AI inventory management tool that was "not meant to replace people; it's meant to augment."

Lum noted that the airline runs 60 flights a day, with 330 days per year to sell a flight at 26 different fare classes, via 15 points of sale, and through 20 major channels, resulting in roughly 154 million possible data points for eight inventory analysts and between eight and 10 pricing analysts.

"There are multiple different price points for each time of year, or each day, and there's no way a human can do this on a daily basis," he said.

The key, Lum said, is to give analysts the context of the data that exists to create the model. "They may not know what the function is or the model underneath," he said. "You give them the contextual data and give them the output for them to trust the model. This has landed very successfully."

He said his team then turned the formula inward, as part of their offensive strategy, to analyse their own data and figure out what the revenue benefit of the data program would be.

"Across the entire network, initially [the program earned] about 1.4% per ask, or £27 million [$35 million] revenue. That is a very compelling business case to the CEO. Even if you're testing and you only take a sample, we're still talking about a £2 million incremental revenue benefit in 2019. That's my budget, so we push all the data engineering needs and funding the data governance processes into the backend," Lum said. "Essentially what we've done is to create, first and foremost, that offensive strategy to fund the defensive side. Because from the offensive side, we realized that there is a whole lot more money to be had with one algorithm than there was on the defensive side in terms of continuing to help with regulation."

Lum said Virgin Atlantic's data team works very closely with the IT department, and there is a data science community within the organization that crosses departmental lines.

"The data science team works with my data product team in order to build the platforms that are necessary for operations folks to bring [their goals] to bear," he revealed. "There are fuel consumption and operational algorithms that utilize the data from the commercial side to ensure that if we have a very profitable flight and a very important set of customers on board, that feeds in real time into our pilots and their tablets today, to make sure they hit what we call a zero, which is arriving bang on time, not a minute later, so that folks can make their connections." ■

Experts Warn on Data Quality for Artificial Intelligence Projects

Data governance strategies that rationalize data, prioritize lineage, and put a premium on quality are necessary before moving into artificial intelligence (AI) and machine learning, according to data chiefs. Emliia David reports from the Buy-Side Technology North American Summit.

Speakers at the Buy-Side Technology North American Summit, held in New York on October 8, said that data continues to be the important missing piece people forget before undertaking AI projects.

For Julia Bardmesser, senior vice president and head of data at Voya Investment Management, AI projects are useless without an understanding of the data fed to them.

"Before you use AI, you need to get your data together. It's cool, but you need to have the basic understanding of the data you have, know what data is good and what it looks like," Bardmesser said. "You can run your machine learning on any pretty much any data you can source, but how much are you going to trust the results if you don't know where it came from?"

Many of the speakers on the panel are already working on AI and machine-learning programs, and have encountered the pitfalls of a lack of quality data and the lack of stringent data management. As a result, they have undertaken projects around this, they said.

"What I've heard is that if you have a lot of data, quality is not important but that's just not true," Bardmesser told delegates. Indeed, that can even complicate the issue further, according to Phillip Dundas, head of technology and change at Schroders. Quality is paramount, he explained.

"We're going through the process of rationalizing the number of data masters or data sources we have, because we've seen that exact problem," Dundas said. "People wanted their data close to them, but what that ends up producing [is] inconsistent results. That's the last thing you want when dealing with clients, presenting one dataset that looks like this and another that looks different."

> "Before you use AI, you need to get your data together. You can run your machine learning on any pretty much any data you can source, but how much are you going to trust the results if you don't know where it came from?" **Julia Bardmesser, Voya Investment Management**

He added the company has a single master list for data around clients or investment within the entire firm.

All speakers noted AI and machine learning have great benefits for their firms despite data issues to contend with.

"Machine learning is the killer app that will ultimately take us on the operational side towards that final nirvana of eliminating the last mile of manual processing," said Michael McGovern, enterprise chief information officer for Brown Brothers Harriman. "The tools are great, but it's the expertise that is important, so we all need to become citizen data scientists."

Voya's Bardmesser noted that putting bots to work in automating operations processes has helped make operations more attractive to employees. She said talent retention increased once the most manual and tedious processes were eliminated.

Bardmesser and the other panelists, however, warned that it is not just data that's important for AI and automation—the ability to optimize processes so as not to automate something that is inefficient in the first place is another necessary step. ■

For Buy Side, ESG, Alt Data Are Necessary Cost to 'Recapture Alpha'

Alternative and ESG datasets hold the promise of delivering better and more predictable returns for investors, but are some firms underestimating the amount of work required to integrate these into their strategies? Max Bowie reports.

Investment firms are increasingly seeking out sources of ESG (environmental, social and governance) and alternative data to deliver alpha, despite the associated costs and practical challenges of acquiring, managing and integrating the data into their investment strategies, according to a panel at the North American Buy-Side Technology Summit, hosted by *WatersTechnology*.

"By using alternative data, you can get so much more [insight]… than the limited information that a company makes available in its financial statements," said Mike Chen, senior portfolio manager at PanAgora Asset Management. "If others are using it and you're not, then… you're being left behind."

According to an audience poll conducted at the event, 50 percent of firms in attendance are already using alternative data to support investment strategies, and another 30 percent are planning to incorporate it into their strategies.

Portfolio managers are moving away from traditional styles of investing and are looking at how to "recapture that alpha," said Lisa Conner, head of client services for North America at Rimes. "We're seeing a lot of questions from our clients about ESG."

This reflects a shift in the composition of the economy worldwide, moving away from manufacturing stocks to where the largest US stocks include Google, Facebook and Amazon, tech companies whose assets are more IP-based and less tangible than a company manufacturing

Dessa Glasser
Financial Risk Group

products. To measure the performance of these companies, one must look at how efficiently a company is run—and ESG factors such as pollution add value to that assessment, Chen said.

Other panelists noted that companies with better ESG practices tend to be less volatile than their rivals, while Dessa Glasser, principal of the Financial Risk Group, who moderated the panel, noted that ESG-focused companies also tend to manage their risk better and be better managed overall.

However, Chen warned that firms shouldn't try to make their strategy fit the latest 'cool' datasets, but should take advantage of them only where they fit a specific investment need.

"For example, if we have a pharma strategy… we come at it as an investment question first, then look for datasets," he said. "We try to think like an executive in an industry, and we ask 'How would that help me?'—then we can see where that data would be useful to us."

Then, Panagora evaluates any ESG dataset in the same way that it would treat any other alpha factor.

"It goes through rigorous testing, just like any other type of research. One factor we assess is 'materiality'—whether an ESG factor is material to a company's core operations. For example, carbon emissions for industrials or material manufacturers. And we find that if a company improves its carbon footprint, its overall performance tends to improve, too," Chen said.

Refinitiv (the former Financial & Risk unit of Thomson Reuters), which aggregates sources of alternative data, also assesses each dataset. "We look at how strong the signal is, and what is the duration of the signal—for example, sentiment for day trading, or something that can give insight over a calendar quarter, or longer term," said Manesh Narayan, head of portfolio management and research at Refinitiv.

But obtaining data in a form suitable to conduct that research and testing can be a challenge in itself, including the application of machine learning and artificial intelligence tools, Narayan noted.

"Machine learning is the base layer of what we're doing—it doesn't matter to customers. But what does matter to customers is the accuracy of the system," said Nathaniel Storch, CEO and co-founder of New York-based natural-language processing and sentiment analysis provider Amenity Analytics. "Some platforms top out at 70 percent accuracy. You wouldn't hire an analyst who was only 70 percent accurate. So we're geared to delivering human-like levels of accuracy."

However, managing these data types with the same levels of precision as mainstream market data comes at a cost: "The cost of bringing in data has gone down, in line with the costs of technology, but the legal and contracting costs have not gone down," Narayan said. "Yes, the cost of ingesting data has gone down alongside the cost of technology, but… it's a 'heavy lift' to hire the data scientists to handle this data," Conner added. ■

SEMANTICS' SEMINAL MOMENT

Data experts testify that right now, the financial services industry is uniquely positioned for semantics breakthroughs that will revolutionize the way data is managed, leading to unprecedented payoffs. Jamie Hyman investigates why, after years of work toward an industry-wide onotolgy, the people passionate about semantics and standards are so excited about the current collaborations.

Financial services is an industry that reveres its watershed moments. From the Great Depression to the dot-com bubble to the financial crisis of 2008, turning points are endlessly analyzed, exploited for value and eventually, the aftermath is put under the same microscope. The scrutiny is for good reason—financial watersheds don't just change the marketplace, they change the world.

However, some pivotal turning points are quieter, subtle shifts that might not seem momentous, at first, but have the power to transform how the market operates. Right now, semantics are having a seminal moment.

"I've never seen anything like it," says Mike Atkin, strategic advisor for the EDM Council, when asked whether an industry-wide semantic ontology is within reach.

David Saul, senior vice president and chief scientist at State Street, has been working on semantics, standards and ontologies for six years. He says the industry is finally at a point where the need for data harmonization is understood and managed by custodians who grasp both the risks of fractured data and the potential for semantic data standards and ontologies to prevent another financial crisis.

"The technology's matured, the standards organizations are working together, collaborating at a global level," Saul says. "This convergence is really all coming together and I'm very optimistic that we're going to see an accelerated update in adoption of semantic data standards that are going to benefit everybody."

> "The technology's matured, the standards organizations are working together, collaborating at a global level. This convergence is really all coming together and I'm very optimistic that we're going to see an accelerated update in adoption of semantic data standards that are going to benefit everybody." **David Saul, State Street**

Saul says the financial industry continually creates increasingly complex ways of dealing with itself and the only way things will get better is through ontologies.

"The problem is real and getting realer," Atkin says. "The ontologies are maturing and approaching material completion. The tools, the vendors that support this are expanding and growing up in capability. Initial use cases have all worked as advertised. Put that together and you've got a perfect storm: requirement—both regulatory and business—capability and maturity all coming together. And it fixes the problem."

The Data Problem

Specifically, that problem is what Atkin calls "the fragmented IT reality," wherein the underlying technology that drives the financial industry has been built up and is managed in silos, which results in repositories of data aligned vertically to their applications.

"So things from data get transformed to meet the data models and software objectives of the systems—a lot of that proprietary—and all of that is designed to drive applications and we manage that vertically. If you multiply that by hundreds, thousands, ten thousands and then across the industry, hundreds of thousands of those things, you have this big transformation problem, where data that represents something precise and real, like an obligation of contract, has been transformed and renamed. Same word, different thing; same thing, different word. We have front-office systems that are fairly straightforward, that don't contain the nuance needed for back, and that is rampant across our industry," he says. "And that is the data problem."

According to Atkin, the solution is a model of reality where data's labels are aligned to their precise meaning via standards.

"The model reality, that's an ontology, the reality of financial instruments and how pricing and processes work," he says. "The standards allow us to assign meaning to the data and process it so that it's not restricted by tables and rows and columns and relational environments that are location-based problems. It's hard to unravel all those joins and builds and locations of all the data because it's all over the place."

Georgia Prothero, principal data modeler at Schroders, says semantics define the real world at a conceptual level and "an ontology is a statement about sets of things and how they interrelate."

According to Saul, a data ontology builds on things that have been around for a while, like taxonomies or data dictionaries, but as a complete representation of not only what the data means, but how it relates to other data.

"For example, we can look at a financial instrument and see the shared characteristics and shared identifiers with the various kinds of financial instruments, whether they be equities, bonds or whatever, and having this degree of mathematical precision describing the data and carrying that along with the data means that when I move data internally within our company from one business area to another, or to a client, or from a client or to a regulator, they know exactly what the data means, what it represents," he says. "It increases the level of data quality to the point that trust increases, and so you can now be sure that this particular transaction, you know exactly what's going to happen with it."

Cross-Standards Collaboration

There are a number of organizations working on developing various standards models, and harmonizing them with others: International Swaps and Derivatives Association (ISDA) is preparing a common domain model for derivatives, FIX Trading Community promotes the use of standards—including the FIX Protocol, a messaging standard for securities—to address market and regulatory issues, XBRL communicates between business systems, and Inter-bank Market Information Exchange (IMIX) is a financial standard for Chinese foreign exchange (FX) trades, just to name a few.

Banks are getting in on the progress, working on solving the data problem through development of internal standards. Schroders' Prothero says the asset management company recognized the importance of a

semantic ontology and established an internal ontology independent of any of their physical system implementations.

"Our semantic ontology is represented by a group-wide logical data model and an accompanying glossary," she says. "Having an agreed internal language has made it much easier for people to communicate business scenarios, and for IT systems to 'speak' to each other."

Among the best and most prominent examples of breakthroughs and collaboration in terms of advancing standards and ontologies, however, are Financial Industry Business Ontology (FIBO) and ISO 20022.

FIBO is a data harmonization standard developed by the EDM Council in partnership with the Object Management Group (OMG) using RDF/OWL (Resource Description Framework/Web Ontology Language).

"We recognize in our industry that the thing that we were focused in on was the reference data about financial instruments and financial process, corporate actions and issuance and transaction pricing and things of that nature. So FIBO is the ontology for that. It's the ontology of contracts," Atkin says, adding that contractual obligation is a "big section" of the industry. FIBO uses XML schemas, and "much of the reference data contractual world has been defined already as part of those XML schemas. So what we're doing is taking the knowledge that exists in those schemas and expressing it as both an ontology in the web ontology language standards. So it is aligned with all of these existing things because all of them are based on contractual requirement."

He calls the work of expressing all of the "very difficult contractual obligations" in an ontology the "underlying foundational segment," and says it is complete, as in, fully aligned with XML schemas.

"[The ontology] is all standardized, it's all been released as standards though our partnership with OMG, so all reviewed, validated, tested and we have a high degree of confidence. Then we have domains. Equities, bonds, mortgages, interest rate swaps and derivatives to some degree—those are all complete, tested and we have a high degree of confidence and they're being used," Atkin says. "It is not a complete ontology, it's not finished, but it is pretty damn good. And if you look at what we currently do, which is mostly communication types of things, trade confirmation types of things, and we use all these XML schemas for that, much of the reference data contractual world has been defined already as part of those XML schemas. So what we're doing is taking the knowledge that exists in those schemas and expressing it as an ontology in the web ontology language standards. So it is aligned with all of these existing things because all of them are based on contractual requirement."

Now, he says, the work is on development projects in RDF/OWL format, so far untested for conformance and in need of further small and medium-sized enterprises (SME) reviews.

"Some of them are really good. Some of them are fairly good. Some of them are just conceptually good. And that's the work that has to be done to materially complete," Atkin says. So now that the contractual concepts are defined, the debate is focused on how best to structure them, and how they relate to each other. "We're a consensus organization. If anyone finds anything that they think is wrong with the way we've captured the ontology, they are encouraged to challenge what's been done. We'll open up a ticket audit, we'll convene a discussion, we'll explore. We get things wrong all the time."

The Big Breakthrough

As part of this collaborative process, the EDM Council is an active voting member and participant in the TC 68 working group "designed to bring all of these niche standards together to create a semantic model for the industry," Atkin says.

Mike Atkin
EDM Council

TC 68 is one of several hundred technical committees formed under the International Organization for Standardization. The first committee, TC 1, was formed in 1947 to standardize screw threads, and active committees focus on developing standards for everything from toy safety to fine ceramics to internal combustion engines. TC 68 covers financial services through the work of three subcommittees: reference data standards, information exchange and information security. ISO 20022 is one of the standards authored, supported and maintained by TC 68.

ISO 20022 was first published in 2004 as a mixture of international standards and went back into revision "to progress the work on technical specifications so that they could be reissued as international standards," says Karla McKenna, director of market practice and standards for Citi Markets and Securities Services, and chair of ISO/TC 68. "Since then, it's come up for revision again. The current revision that's going on is adding semantic capability to ISO 20022."

ISO 20022's core messaging syntax is XML, so it's extensible, as opposed to fixed-length code qualifier tag pairs found in past transactional and messaging standards. "It takes it up a notch as far as being able to be adaptable. It is a recipe for putting concepts into a common repository. You can use the same recipe for more than one business domain," McKenna says.

Jim Northey
FIX Trading Community

McKenna describes a "shift in focus" among the committee as it "worked with the standard more and observed the industry around us. This is where the motivation for the introduction of semantic capability into ISO 20022 came from, where the emphasis is now on interoperability with other standards. There are other standards that have grown up within financial services, and we were looking

through semantics to be able to make sure that they interoperated or related to or could leverage in some way ISO 20022. This is where the kernel of the idea and motivation came from."

Initially, the committee attempted to harmonize and incorporate different standards' concepts into ISO 20022, but now they are building a semantic representation of the ISO 20022 metamodel.

"The current effort is not focusing on trying to be able to map at the business domain level, the business concept level, but more at the semantic and format level," McKenna says. "The ISO working group has created a tool to normalize other standards into a common semantic format that is derived from the ISO 20022 metamodel itself. So it shows us the equivalent concepts of ISO 20022 in other standards, but also, because these standards have grown up to be able to execute and serve different purposes, it also shows us enriched concepts that they carry in the standard."

This cross-standards work is not a mere kumbaya moment for data professionals: It is a core component of TC 68's strategy to advance and enrich the standard.

Jim Northey is a technical committee co-chair for the FIX Trading Community and will replace McKenna as chair of TC 68 this coming January, once her term expires. He says TC 68 "had a bit of a breakthrough" when investigating whether semantics would help map ISO 20022 across different standards, resulting in the shift in focus McKenna described, which led to work toward a semantic version of the ISO 20022 structure, executed so that other standards could be represented—as they currently exist—in that format.

"We're not expecting any of these other standards to change their formats," Northey says. "All we're asking is give us a translation into this independent semantic format and then we can start to use all of the semantic tools for querying, knowledge discovery, automated machine learning, against the standard itself to start to identify new facts and new knowledge."

Georgia Prothero
Schroders

Karla McKenna
Citi

He calls it "comparing apples to apples" and says if the plan sounds simple, that's because it is. "It wasn't so obvious until we thought of it and then, like many good ideas, seemed obvious."

Before the breakthrough, Northey says, FIX spent years and "quite a bit of money" attempting to map the FIX standard into the ISO 20022 model, a process he says was "very expensive and frustrating because things didn't quite align, and once it was mapped, it wasn't really useable by any practitioner."

The committee found that when taking a semantic approach, the representations are "more extensible and powerful tools where we can actually start to look at and compare and start to use them," he says, adding that TC 68 plans to bring in other reference data standards. "We have a number of other industry standards we want to get represented in this format so that then we can start to do work in terms of interoperability, improving processing."

No VHS vs. Beta

A combination of bank needs, regulatory requirements, industry collaboration and ontology creation and implementation is what will move semantics toward becoming a workable, robust solution, says EDM Council's Atkin.

"A lot of people are working on this," he says. "If you look at any bank, most of them participate in some form of conceptual modeling as part of their processes. Conceptual modeling just means, do we understand what these things are, how they work, can we model them so we can implement them in our environments? [FIBO] is a standard conceptual model. A lot of people are working on the same problem, and now we're just trying to collaborate and bring it together into a single view."

Schroder's Prothero underscores Atkin's view of how the industry will collaborate, noting that Schroders' glossary "has been cross-checked to FIBO and wherever there is a match the FIBO definition is used and a citation is made to the FIBO term."

She predicts the first adopters of an industry-wide semantic ontology will be vendors, "ahead of take-up by bigger organizations."

The way Atkin sees it, current initiatives fall roughly into three camps. First, regulators collaborating with each other and issuing requirements documents, essentially stating, "here's what we regulators need from the industry so we can do our job, and we need it to be consistent, so we're going to specify what that is." The second camp is a charge led by the Financial Conduct Authority (FCA) toward smarter regulatory reporting through technology, and their "better way" matches the views of the EDM Council: an ontology "with regulatory business rules, specified in machine-executable language, so that there's no confusion about meeting, and we can all adopt standards and move to the next stage," he says. The third camp is banks saddled with "environments that do not allow them to integrate, link and leverage" their own capabilities, and they want to fix that. Both the regulator and bank problems can be solved the same way, Atkin says, with "ontology plus standards plus executable business rules, and that's the direction we're all covering, moving in."

In the meantime, he adds, "it doesn't have to be perfect to be functional."

It's an unprecedented level of collaboration, and State Street's Saul says it's just the beginning.

"I'm very optimistic that we're going to see more collaboration rather than splintering into multiple competing standards," he says. "This is not going to be, in my opinion, VHS and Beta. It's going to be more like the DVD standard where all of the industry converged on one." ■

Alt Data's Ethical Day of Reckoning

Alternative data, an industry projected to be worth over $350 million by 2020, is no longer the hedge fund industry's best kept secret. Amelia Axelsen investigates whether financial services is on the brink of its own Cambridge Analytica moment or if it is simply time for an alt data ethics evaluation.

In April 2017, during an atypical criminal case fitting for an era driven by technological advancements, a jury convicted Nan Huang, an experienced data analyst at Capital One Financial Corp., of insider trading.

Huang, hired to investigate fraudulent credit card activity, created detailed spreadsheets of private consumer credit card purchases from Capital One customers, amounting to 2.4 percent of the company's sales. He then cross-referenced the non-public information with historical, publicly available data to analyze and create revenue projections for 226 publicly traded retail companies, which he used to execute securities trades before companies' quarterly earnings announcements.

Using statistical analysis of the combination of the data, Huang traded in 105 out of the 226 companies amassing a total of $4,403,545 in positive profits, or a 12 percent return on his initial investment, according to the Securities and Exchange Commission (SEC).

Stephen Graham, an SEC employee with expertise in economic and statistical analysis, testified during the case and told the jury the consumer credit card data used was a demonstrably valuable tool that helped Huang make profitable trading decisions. Although found guilty on insider trading charges, Huang faced

no legal repercussions for exploiting confidential, individual transaction data for monetary gain.

SEC v. Huang is the first case evaluated within the confines of the law to determine financial executives' impropriety when using alt data, and currently, no case law exists on the use of alt data following the implementation of Europe's sweeping General Data Protection Regulation (GDPR), enacted to protect individual data privacy. Although Huang violated Capital One's employee policies by obtaining the data without company consent, consumer credit card data sold to third parties remains one of the most lucrative sources of data with 38 percent of investment funds using it for an edge in the market. Web data was the highest at 48 percent, followed by social media and sentiment data at 36 percent, satellite at 29 percent, and geolocation at 24 percent respectively.

The legal ramifications of insider trading in *SEC v. Huang* were clear as the verdict was announced, but the ethical implications of using personal information, such as consumer purchases, are the responsibility of each individual firm. As the scope of personal data collection has become etched into the public consciousness—largely due to Facebook's breach of trust in the harvesting of personal information from its website for a British political consulting firm, Cambridge Analytica—firms using personal information such as credit card and social media data are on track to get swept up in the backlash.

Diana Ascher, who received her PhD in information studies with a focus on data ethics at the University of California, Los Angeles (UCLA), began her career in business writing at Bloomberg. As the co-founder of the Information and Ethics Institute and the director of Information Studies Research Labs at UCLA, Ascher is well-versed in data ethics, and she predicts that fallout from data breaches in the technology sector will lead to an examination of alt data practices in financial services and particularly, fintech.

"If firms look longer term, they will find it is in their best interest to consider data ethics at the outset and not in retrospect," says Ascher. "Ethical data practices will help their decision-making in terms of what data is collected, stored, and transmitted, and how we think about data, the consumer, and the ripple effects of sharing even pseudo-anonymous data with third parties."

Alt Data Anonymity

At a Reuters Newsmaker event in London in July, Financial Conduct Authority (FCA) chair Charles Randell painted a stark picture of a world where huge quantities of data gathered from every aspect of our lives is owned by few, resulting in a population that is "prisoners to technology." Although GDPR has fundamentally altered data collection practices in Europe, Randell called for a continuous debate on how to ensure innovation in financial services, from algorithms and big data, remains a "force for good" and doesn't reduce people to numbers.

"We need to anticipate the fundamental questions that big data, artificial intelligence and behavioral science present, and make sure that we innovate ethically to shape the answers," he said.

> "Selling the data that will never go away. All this goes to business models that are tailored by apps, publishers, and the industry. I agree with critics that apps and companies have to be better at showcasing and being concerned with the users and how the data is being used."
> **Thomas Walle, Unacast**

David Saul
State Street

Data vendors and global companies have more access to a plethora of personal data than ever before. Datasets are being compiled on where we shop, how we eat, our political preferences, insights into intimate relationships, browsing history, and even our location, among many other facets of our daily lives and activities.

Individual consumers using an app may not be aware that their location is being tracked to produce data on how many times they visited a Walmart parking lot or why that's important for statistical analysis, but hedge funds eager to capitalize on this information are buying in. According to Alternativedata.org, a website run by former buy-side and sell-side data analysts, alternative data providers have grown from a little over a 100 companies in 2010 to over 350 with total buy-side spending on alt datasets expected to exceed 1.7 billion dollars in 2020.

In order to be compliant with GDPR, vendors and financial companies are required to make datasets anonymous. Anonymizing datasets is touted by vendors as not only a way to avoid legal repercussions, but as a crucial component of ensuring consumer privacy.

Managing director of enterprise at Refinitiv (formerly Thomson Reuters) Marion Leslie says the use of personal data does not focus on individual activity, but rather, produces a macro view of behaviors that then provide a landscape view of market movers.

"The part of the industry that we serve and the services we provide are absolutely not about individuals, it's about market performance, prediction, and insight into how we believe or insights into how the customers can used to predict how markets behave. Certainly for us there is absolutely no interest in individual data," she says.

Ensuring Data Protection

Winston Maxwell, a corporate partner at Hogan Lovells international law firm who specializes in media, communication, and data protection laws and assists clients with global data privacy governance programs and implementation of compliance, says financial firms need to be responsible for their agreements with alt data vendors. Maxwell says it's essential for firms to conduct proper due diligence checks with each vendor before using their alt datasets.

"Investment advisors are very excited about alternative data, but the sort of training issues that we've been focusing on with financial institutions related to GDPR is ensuring in [firms'] contracts with the vendor that you have appropriate reps and warranties to make sure the data that's being provided to the extent is fully anonymized, comes from a legitimate source, and has either the permission of the data subjects or some other legal basis to be used," he says.

Maxwell notes that methods to gather and sell data are not always black and white, and that anonymization is one of GDPR's gray areas. When referring to anonymous data, he said, "it gets pretty tricky. You get into the weeds of European case law and GDPR interpretations and what it means to be anonymous." Hee adds, "What's even worse is that what's anonymous today may not be anonymous tomorrow because you may have tools or artificial intelligence and machine learning that can take a dataset and make it not anonymous."

Under GDPR, data with individual identifiers removed is considered anonymous and not subject to the same protection requirements as data where the subjects are easily identified. But the wording of the second part of GDPR Recital 26, related to anonymous data, makes the law a bit tricky.

"To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments," GDPR Recital 26 states.

In a study by 1 Media Lab at the Massachusetts Institute of Technology (MIT) researchers looked at the re-identifiability of three months of credit card records for more than 1.1 million individuals' credit card metadata and concluded that anonymized financial metadata can be re-identified with spatiotemporal information. For instance, simply knowing the approximate price of someone's coffee even when other key identifiers are removed or scrambled increases the chance of someone being identified.

Researchers took it a step further and studied coarsened data, which consists of data recoding and variable suppression, and deemed that coarse data still provides "little anonymity."

Diana Ascher
UCLA

Ascher notes that the industry has a problem with anonymity and that it should no longer promote datasets as anonymous even if data is encrypted and aggregated.

"Vendors are looking at short-term gains, not long-term ethics," she said. "There isn't such thing as secure data, even if we have the strictest protocols in place to protect data today, we have no idea what type of technologies are going to be developed tomorrow that can overwhelm those systems or work around them."

Alt data provider Unacast, a geolocation data provider seeking to "understand human mobility," complies its data from apps and location data companies. Unacast recently raised $17.5 million in funding with significant investments from WhiteStar Capital, Open Ocean Capital, and European telecom giant Telia. Norwegian native and Unacast's co-founder and CEO, Thomas Walle, says many geolocation vendors from Europe fled to the US due to GDPR concerns.

He mentions an increase in the number of hedge funds and asset managers interested in the datasets Unacast sells, which Walle says helps deliver insights on where people work, live, what types of stores they visit and at what time. He notes that it is possible to link anonymous data back to individuals, but asks, "Do companies really want to do that? Do they benefit from that? No." He adds that there are very few instances of data misuse in the industry.

Walle says a primary driver for firms fleeing to the US, rather than complying with GDPR mandates, is the inability for certain vendors to account for all the personal data used when individuals decide to opt out. The opt-out function requires any data collectors to delete all the data associated with specific individuals

if those individuals choose to limit the sharing of their data. If a vendor doesn't have the technological frameworks to locate all the data for this requirement, it can be a challenge for compliance, he says.

GDPR is championed as creating a better relationship with users and data collectors through responsible, transparent data practices, Walle says. Although the number of geolocation data providers in Europe has gone down, data quality has increased among those who remain, which is significant for hedge funds, he adds.

After GDPR went live, web users were inundated with user agreements and consent forms, necessary for compliance and designed to give insight into how personal data is used, as well as how to opt out. Still, critics argue that convoluted consent forms don't go far enough to explain how data is being used. Walle says the opt-out rate is "next to nothing," sharing a case about a partner that has 2 billion users, but only about 30 users withdrew their consent.

This could be chalked up to individuals being content with the way their personal data is shared or it may indicate a lack of consumer awareness around the importance of consent forms for data protection. Ascher says it's the latter, noting that the expectation that ordinary consumers know their data is being monetized and sold is problematic.

"Consumers are playing catch-up in terms of information literacy and what's being done with the information that they are transmitting," she says. "People don't realize they're giving away their personal data for free and that data is being monetized by big companies and sent to other interested parties, who not only are making investment decisions based on personal information that's gathered a variety of ways, but also can limit your access to opportunities based on the data that is associated with your profile structurally."

Marion Leslie
Refinitiv

Winston Maxwell
Hogan Lovells international

FCA's Randell similarly stated that in order to engender trust between companies and consumers, companies must articulate in a coherent and easily understood manner a firm's approach to using consumer data.

"By good communication, I don't mean pages and pages of obscure disclosures, disclaimers, and consents. I mean short and readable statements that make it clear what firms will and won't do with their customers' data," Randell said, adding that data statements should be developed in partnership with consumers, "not imposed on them."

Data privacy and sharing practices are magnified by data breaches, such as the October 2018 news about a bug in Google's API for its Google+ service, which allowed third-party developers to access information on users and their friends. The public is no longer idle to data practices, evident by Californians' votes in June 2018 to put data privacy at the forefront of legislation by tightening consumer laws through the California Consumer Privacy Act.

But for hedge funds and financial firms, Ascher asks: "How much data do you really need to make informed decisions that put you ahead of your competitor?"

Getting Ahead with Ethical Frameworks

David Saul, chief data scientist and senior vice president at State Street Corp., regularly assesses and proposes new technologies and says financial firms need to distinguish between good intent or harmless intent and use data appropriately.

"I think we're much better off approaching [data ethics] intelligently and putting appropriate regulations in place before we have a major incident," he says.

For both Saul and Gideon Smith, a 20-year quant veteran at Rosenberg Equities, a subsidiary of AXA Investments, applying thorough vetting processes for alternative data providers has been fundamental for both compliance and ethical governance. Smith says both vendors and datasets are evaluated for quality, reliability, accuracy, sourcing, and ethics.

"If companies are transparent about how they use the data and consumers and companies see that very clearly, it's not only going to protect the data, I believe it's going to generate greater business, because when people have more trust in the data then they're going to invest more," says Saul.

For Ascher, avoiding a data scandal means staying ahead of regulations and technology. She suggests requiring certification programs for data ethics and applying a code of ethics for data management would be steps in the right direction.

"Many companies have no data management plan and have no transparent communication of what that ethical code is. As consumers become more literate about these issues, especially in light of lawsuits against Facebook, Twitter, and social media platforms, we really need folks to be talking about other sectors like fintech, where financial information is tied directly to personal information and it is very valuable," says Ascher, adding that if consumers refuse to give up their data for free, it makes it harder for companies to capitalize on individuals' personal information.

"Selling the data, that will never go away," says Walle. "All this goes to business models that are tailored by apps, publishers, and the industry. I agree with critics that apps and companies have to be better at showcasing and being concerned with the users and how the data is being used." ■

Data-Driven Regulators: Handling the Uptick in Regulatory Reporting

With the abundance of existing and new reporting obligations to comply with, how can regulators hope to be able to handle the volumes of data they now receive? Wei-Shen Wong learns what some regulators in Asia are doing with the data they collect.

The Loch Ness Monster, the Bermuda Triangle, regulators that actually do anything with all the data they now demand from financial firms: all myths, or is there an element of truth? Do regulators really collect massive amounts of new data, rubber-stamp them, then lock them away deep in a JRR Tolkien-inspired mountain dungeon, never to be seen again?

Firms grappling with the challenges of collecting and submitting the volumes of data mandated by regulators might be forgiven for believing that managing those submissions from all firms would be all but impossible. Regulators, though, are keen to dispel this myth.

Speaking at the European Financial Information Summit (EFIS) in September, Olga Petrenko, leader of the market data policy team within pan-European regulator the European Securities and Market Authority's (Esma's) markets department, said that "the data nowadays is being used across multiple areas and domains of Esma's activities. Supervision is obviously a part of it—supervisory convergence—and equally all the financial stability and financial market development reports are being done on that basis."

Compared to Europe, central banks and regulators in Asia each have their own individual reporting standards—though there are

similarities in the reports sought by all regulators, particularly around investor protection, says Neil Thomas, head of sales for Asia at SIX. "In the EU, the regulators can leverage a standardized reporting requirement, which does have certain benefits. However, there is no unified format that regulators in Asia use; it's primarily in their individual format and platforms," he says.

Laurence Van Der Loo, a senior manager in the CEO's office at the Asia Securities Industry & Financial Markets Association (Asifma), says in some instances, regulators asking for the same information multiple times leaves the impression with the industry that either they are not able to cope with the sheer volume of data, or they are not equipped to use it for multiple purposes.

"There is also the impression among market participants that regulators do not always know exactly what data to ask for, which leads some regulators to apply a 'shotgun' approach—i.e., asking for a very wide scope of data in the hope that what they need will be included in the reported data. But then they don't use it anyway," she adds.

However, Van Der Loo notes that many regulators in the region are working to improve on this, such as the Monetary Authority of Singapore (MAS), which has vowed that financial institutions will not need to submit the same data twice.

Minimizing Data Duplication

A senior executive responsible for global markets at a Japanese securities firm says one of the issues facing financial firms in Japan is the number of regulatory reports required. "In Tokyo, there are many regulators. Depending on the type of firm and transactions to report, it could range from between 400 to 800 regulatory reports. But one report is submitted to 10 to 15 regulators, so in total, we have a lot," he says, adding that sometimes certain regulators are not using the reports to support their core function.

"They don't raise any questions and they're not using our reports for the purpose of analysis," he says. "So the effort level is quite high. We need to send reports by the target time, and if we make a mistake we may need to resubmit them, with an apology letter at times."

Regulators in Japan could not be reached for comment.

In March, MAS revealed its roadmap to transform its approach to data collection from financial institutions. Not only will this reduce the amount of resources and preparation time needed to comply with requests by MAS; it will also make it more efficient for the regulator to process and analyze the data collected.

"MAS is transforming the way we collect data from financial institutions (FIs), to reduce duplication of data and automate data submission by FIs," a MAS spokesperson says. "MAS will reuse the data the FIs provide in their regulatory submissions, and ensure that FIs need not submit the same data to MAS twice."

In addition, all financial institutions operating in Singapore are required to submit all regulatory returns to MAS in machine-readable formats to make the data collection process more efficient and to minimize the risk of human error. From 2019, all responses to new surveys and ad-hoc data requests must also be provided in machine-readable format.

As for data that can be aggregated in different ways, the central bank is working with the financial industry to collect more granular data on underlying transactions rather than aggregate statistics. According to MAS, this more detailed approach reduces the reporting burden on institutions and allows for better identification of potential systemic risks.

Alexander Kling
Synechron

The regulator revised some of its reporting standards for banks in May to account for these changes to data collection. For example, institutions can decline any request from MAS for structured data they have previously provided. The aim is to eliminate all duplication in data submissions by the end of 2019.

The changes include collecting more detailed data on banks' assets and liabilities by currency, country and industry. "Greater granularity allows better identification of potential risks to the banking system," MAS officials said in an earlier announcement.

This approach is exemplified in its revisions to MAS 610 and MAS 1003, which govern the submission of statistics and returns by banks. The number of data points that banks must submit to MAS has increased to about 340,000, compared to 4,000 data points previously, as a result of more granular data fields required.

MAS is unique, because it is not only Singapore's central bank, but also its financial regulatory authority. According to MAS' website, it administers the various statutes pertaining to money, banking, insurance, securities and the financial sector in general. This differs from other countries in Asia, such as Japan, where firms must deal with different regulators covering various sectors.

Hong Kong's Financial Services and Treasury Bureau and HKMA—which does not regulate financial firms—declined requests for comments.

Data Makes the World Go Round

Alexander Kling, director at technology consultancy Synechron, says data collection has become increasingly important, especially since the global financial crisis. Many

post-crisis regulatory reforms spawned the introduction of new guidelines, such as Basel III, IFRS 9, Dodd–Frank, EMIR, Mifir and Mifid II.

"Regulators around the globe have moved toward both increased frequency and granularity of data collection points for the purpose of reporting and predictive analysis at both the macro and micro level," Kling says.

He sees regulators establishing targeted initiatives to cope with vast volumes of data coming their way.

"Early last year, the MAS introduced a Data Analytics Group, aimed at unlocking insights, enhancing supervision and making regulatory compliance more efficient for financial institutions. However, this discipline is still nascent within regulators and central banks, with varying degrees of maturity around sophistication and tooling," he says.

The DAG will comprises three units: the data governance and architecture office; the specialist analytics and visualization office (SAV); and supervisory technology office (SupTech). The DGA develops data management policies, manages data collection and quality, and maintains the MAS' data catalogue and publish the central bank's official statistics.

SAV, together with other departments within MAS, will conduct data analyses and help to improve data capabilities using reusable tools and code libraries. It will work with MAS' IT department to design and implement technical infrastructure to support MAS' data analytics work.

While the amount of data has increased exponentially in the last two years alone, Alexander Dorfmann, senior product manager at SIX, says there are new possibilities for both businesses and regulators.

"Regulators are implementing these measures and controls as a mechanism to ensure compliance and adherence to a code of conduct. If a company is then perceived to be in breach, they will use the collected data to look for evidence of a breach. Even if it may seem that the entire industry, including regulators, is over-burdened by the data volume, we recognize that regulators are increasingly using this data to evidence a breach as we have seen in the cases of anti-money laundering (AML) or sanctioned securities," he says.

Alexander Dorfmann
SIX

Singapore's MAS is not the only regulator in the region making strides in this area. While the Australian Prudential Regulation Authority (APRA) declined to comment for this article, it is currently undertaking a major data modernization project to replace its Direct to APRA (D2A) data reporting tool.

APRA collects, stores and analyzes data to inform its core supervisory work, assist with policy development, and support other government agencies, including the Australian Bureau of Statistics (ABS), Australian Securities and Investments Commission (ASIC), and the Reserve Bank of Australia (RBA).

According to APRA's website, the data transformation program will fundamentally change all aspects of APRA's data management—how it collects, stores, and analyzes data, as well as how it can use data to innovate.

The new solution will enable data uploads, so it is less dependent on manual data entry, compared to D2A. It will also be web-based, therefore simplifying maintenance, and will be able to adapt as reporting requirements and technology continue to evolve.

APRA has already conducted an industry-wide engagement and published the findings in July 2018. Four key themes that emerged are that the solution should have multiple submission channels and formats, a secure web-based portal, an improved data validation and query experience, and test environments for transition and ongoing support.

Currently, entities submit data through D2A using two primary channels—data entry and file upload—either in extensible business reporting language (XBRL) or extensible mark-up language (XML).

In August, APRA issued a request for tender, seeking software companies that can provide a data collection tool that meets its requirements and standards. It intends to finalize the selection of the new data collection solution this year and aims to start the rollout and transition by the first quarter of 2020.

Meanwhile, ASIC would only say that whatever disclosure it requests from companies, managed investment funds and other relevant stakeholders, it must be clear and useful information that is readily digestible and informative. "That remains the goal at all times," an ASIC official says, rejecting the notion that that regulators collect data for no purpose.

Under Pressure

This still means, though, that financial firms must redouble their efforts around collecting, organizing and maintaining their data, and do so in a way that is acceptable to regulators, a responsibility that may reach beyond the scope of where a firm is headquartered.

Synechron's Kling says undoubtedly, the regulatory impact for financial institutions with a global presence is more burdensome, because not only are they required to meet local obligations, but also international requirements. Banks in Singapore with operations or presence in other jurisdictions will also have to comply with additional regulatory regimes mandated by those jurisdictions.

Laurence Van Der Loo
Asifma

This leads to higher compliance and operational costs for banks. Kling says, "Irrespective of the size and extensiveness of the bank's footprint, there is pressure from both financial and operational headcount perspectives for the banks."

He says increasing regulatory requirements have pushed many financial firms to "level up" their IT architecture and infrastructure—for example, the implementation of global data lakes across many banks, which is used as the golden source for regulatory requirements across jurisdictions.

In turn, this raises debate over whether firms should buy or build. Kling says Synechron sees just as many banks building solutions in-house as well as implementing vendor products—though there is also an emphasis on harmonization of data architectures and data sources. This is to centralize efforts, ensure common standards, and remove duplication. "Getting the outputs 'right' is no longer the only priority as regulators are pushing for transparency around defensible, robust, scalable and automated processes to produce the numbers," he says.

No Easy Task

So how can firms achieve the end goal of reducing heavy data processing and automating manual processes? Dorfmann says because the data universe within financial markets is highly fragmented, the exact needs of different types of participants are rarely met without significant effort.

"Having said that, we see that staying focused on the rules imposed by regulators globally will push the financial industry to a new level playing field in order to meet compliance deadlines. Streamlining and aligning the underlying data and content is, and remains, a huge burden. It also raised the question of how to best address these challenges," says Dorfmann.

Olga Petrenko
Esma

Firms are beginning to realize the systems and infrastructures that have kept them running will not suffice to amass and sort through the necessary data.

Modernizing current regulatory reporting practices is not an easy task, says Van Der Loo, because many firms still rely on manual, outdated systems, but regulators can help the industry progress.

"For example, instead of asking for specific reports, one solution could be to collect standardized raw data that can be sliced and diced multiple times, according to the needs of the regulator, while reducing the regulatory burden on the industry. The MAS in Singapore is going down this route," she says.

"When doing this, it is key to use as much as possible existing data standards and dictionaries (e.g., ISO20022). Another good example is the use of LEI for the consistent identification of counterparties."

One result of the improvement efforts on all fronts is an industry becoming more collaborative and eager to agree on new data standards that will provide consistency across numerous disparate regulatory reforms and jurisdictions, Dorfmann adds.

And regulators can promote further automation and the adoption of new regulation-focused technologies by asking for data in a machine-readable format. "This will lead to better fraud detection, supervision and enforcement for regulators, and higher efficiencies and lower costs for the industry," Van Der Loo says.

As demonstrated by MAS and APRA, the capability to handle new levels of reported data requires new technology infrastructures for both firms that submit data and for regulators that receive it. As long as data volumes continue to exceed what humans can physically process, regulators who are not automating data analysis might as well be doing nothing with the data at all. ■

EVOLVE OR DIE:

Asset Managers Cultivate Data Science Teams

Firms are using machine learning and natural-language processing tools to scan an ocean of text and images, as well as big data and proprietary datasets—no longer to grab an edge, but merely to remain competitive, reports *Risk.net*'s Faye Kilburn.

In the time it takes the Earth to rotate about its axis, internet users will generate 2.5 quintillion bytes of new data. That number, a calculation by IBM, is mostly a slag heap of digital dross. But it is a mountain asset managers can no longer afford to ignore. Whether to spin alpha or just survive, asset managers need to separate the meaningful and profitable from the futile and worthless. And if humans can't do it, a robot will.

"Data science, big data and machine learning are all becoming a necessity just to compete," says Anthony Lawler, co-head of GAM Systematic, a $4.7 billion quant fund. GAM has robots working on predicting asset prices and even vetting how fruitful new sources of data will be.

Lawler compares this moment to the early 1980s when a small startup called Bloomberg began selling its now-ubiquitous terminal. Those terminals put a glistening smorgasbord of new information at the fingertips of asset managers.

"If you didn't have Bloomberg, certainly within one year of that coming out, you were at a disadvantage, and you would be behind the price discovery curve by not having that technology," Lawler says. "The current evolution is not dissimilar."

That is, if you're not using available data in real time, other people are. And if you fall behind, "your returns will become less competitive," Lawler says.

Some of the world's largest asset managers are on board.

Goldman Sachs Asset Management has incorporated big data, and by extension, some type of machine learning, into around half of the factors it uses to select stocks (it uses several hundred in total). Pimco uses data science to inform its bids in US Treasury auctions, while a 'muni bot' helps JPMorgan Asset Management to quickly source investments for its municipal bond portfolios.

"Data science, big data and machine learning are all becoming a necessity just to compete."
Anthony Lawler, GAM Systematic

Hunting for Signals in Data

The most daunting task facing data-driven investors may be grappling with the gargantuan volume of data. Some of that is records, trading data and number sets, a vast field now manageable through large-scale computing. The even bigger rest is unstructured data—mostly text- and image-heavy information—that firms are combing for the faintest hint of an investment signal. And according to analysts at Gartner, 80 percent of new data is unstructured.

So GSAM is spending most of its data science budget on natural-language processing, says Nick Chan, a managing director on the quantitative investment strategies team. The firm is using it, for instance, to probe tens of thousands of news articles, analyst research notes, and other text information. The objective is "to pick up on subtle connections that companies have with one another that often go unnoticed by most investors," Chan says.

The firm is also looking at satellite images of parking lots, he says. The breadth of stores and locations, plus foot traffic, combined with online sales can be used to forecast revenues and earnings. GSAM is also looking at average purchases in aggregated credit card data to better gauge earnings and revenue prospects of companies.

The same technology can also make sense of trading data in-house that would otherwise be too messy to process, says Chan. "One of the benefits of being part of a larger firm like Goldman is that we have access to a lot of publicly available information that is often difficult to aggregate or expensive to obtain," he says. "So, for example, like tick-level option data is very dispersed information, it's very unclean information, it's very difficult to get an aggregated centralized source of that data anywhere in the industry."

Goldman is hardly the only firm looking closely at internal data. Pimco, one of the world's largest bond investors, is focusing on building an algorithm to find the optimal way to participate in Treasury auctions.

Mihir Worah, Pimco's chief investment officer for asset allocation and real return, says the firm is scrutinizing data on every Treasury auction going back several years, along with proprietary data on what Pimco did at those auctions, to calculate the optimal bid/ask level, trade size and execution venue for trades going forward.

The firm has also been combing through 20 to 30 years of information on mortgage-backed securities, representing billions of data points. An outside artificial intelligence expert has made its algorithms faster, and the firm can now assess around 20 percent of that data, up from a meager one or two percent. Within six months, Worah says, the firm's prepayment models were 10 percent more accurate than previous ones in predicting near-term outcomes.

In addition, Pimco looks at large dimensional datasets, throwing in thousands of economic variables, and lets a computer figure out what's new. Using around 1,000 different data points—jobless rates, wages and bond yields or what the stock market might be doing in a given country—the machine spits out answers. The firm then compares this to what its fundamental analysts are saying.

The firm is also looking at anonymized payroll-processing data—such as hires, layoffs, the number of jobs created, and salaries across different industries—to try and see economic trends before the rest of the market does.

But even with the advantages of data science, Worah says Pimco has no plans to allow machines to automatically enter orders.

"We are weighting the model projections pretty heavily, but at the end of the day we still have humans getting the input from the model," he says. "We're not letting the machines automate investment decisions. That's not us."

John Chisholm
Acadian Asset Management

JPMorgan Asset Management's data science effort is also focused primarily on using natural-language processing for alpha generation—for instance, using machine-learning algorithms to scan large documents for new information on bond issuances. The firm built a recurrent neural network to determine the context of words, so its robots can distinguish between a reference to a chemical bond and a financial bond.

The firm is also using machine learning to speed up the process of finding investments in illiquid markets, such as municipal bonds.

"A pain point for the traders is sourcing liquidity in the muni market to fill portfolios—it's very fragmented and there are thousands and thousands of Cusips," says Ravit Mandell, chief data scientist in the intelligent digital solutions division at JPMorgan Asset Management.

To help the traders, Mandell's group set up electronic connections to dealers and built a "bot" that scours the market for bonds to fill portfolios. "This provides a huge time-save for our traders, allowing them to better focus their efforts, and also allows us to offer our muni investment products to more parts of the market at scale," she says.

Refining the Risk Profile

JPMorgan is also exploring the application of this technology to risk management problems, like anomaly detection.

"One method of detecting anomalies is using clustering, which is an unsupervised method of machine learning," says Mandell. "This allows us to feed huge datasets into a model. The model then plays back a clustering visual that shows relationships and possible outliers based on the data attributes that were fed in."

Ravit Mandell
JPMorgan Asset Management

Such techniques could help the firm detect valuations errors for complex instruments, she says.

Most asset managers are only just beginning to think about the potential uses of machine learning in risk management.

"Most investors so far haven't focused heavily on those things," says John Chisholm, co-CEO and co-chief investment officer at Acadian Asset Management. "Investors are further behind in forecasting risks than they are in forecasting returns."

Vis Nayar, deputy chief investment officer for equities at HSBC Global Asset Management, also says there's still work to be done

in applying data science and machine learning to risk management and portfolio optimization.

"I think the danger is that we tend to say we've done everything we need to on areas like risk modeling because it's a dry topic," he says. But risk management and optimization "are not done and dusted," he adds, "and some of the technologies that are out there today can allow us to do a better job and need constant revisiting."

One area Nayar is exploring is covariance matrix estimation. Asset managers have never had a perfect view of the exact correlation of volatilities in equities, leaving a lot of decisions to be made under a shroud of uncertainty. HSBC's data scientists are trying to fill the gaps.

"Covariance matrix regularization, probabilistic clustering, network-based machine learning and graph theory algorithms are some examples of techniques which potentially could be utilized towards this direction," Nayar says.

In the past, computers were not sophisticated enough to handle covariance models, so risk overseers were forced to reduce the dimension of the problem. Newer machine-learning techniques can better handle those tasks, he says.

Acadian's Chisholm says machine learning can be used to "vary some of the parameters" of risk models, "looking at whether there are some regime-switching components to the risk model, and what kind of lookback period to use in different environments."

Still, Acadian continues to get most of its added value from more traditional statistical techniques, Chisholm says. But the firm is seeing benefits from machine learning and is committed to it. "We don't think it's a magic bullet," he cautions. "You still need human ideas in terms of the intuition around what the relationships in the market are."

Mihir Worah
Pimco

Will Kinlaw
State Street

A Big, But Green Field

But there are hurdles to overcome. Will Kinlaw, senior managing director and head of State Street's academic affiliate, State Street Associates, says most of these newer datasets do not have that much history because the data has not been around that long. That limits the ability of artificial-intelligence and machine-learning techniques to pick up on patterns and assimilate them.

State Street is currently exploring the nooks and crannies of its businesses to identify datasets that might be valuable to clients. Although the primary focus is how that data could offer a different perspective on investment decision-making, Kinlaw says it's difficult to disentangle it from risk management.

"Most managers are evaluated based on their risk-adjusted returns," he says, "so it's definitely the case that if a portfolio manager can use alternative data to sidestep a drawdown or reduce risk at the right time, that's a driver of performance over the long term."

Kinlaw says machine learning and artificial intelligence are slowly catching on. "We have numerous clients who are using web-scraped media sentiment and inflation signals as inputs to their strategies," he says.

Also raising its sights, GAM Systematic has appointed a data tsar whose assignment is to ferret out traditional and alternative datasets to bring in-house and make sure the fund's database is never compromised by corrupt or inaccurate information.

The asset manager is also using natural-language processing to detect shifts in tone in analyst reports; for instance, whether a given analyst has turned more positive or negative in their view from prior reports—a task no human being could do at that scale, Lawler says.

The firm is even extending its data science program into the process of cleaning data, using machine learning to run analytics on datasets to see how accurate and reliable they are.

"If we decide that it's an interesting dataset, we will often use reasonably sophisticated data science on the dataset to see if we find signal value," Lawler says.

In any case, others are bolstering their own efforts.

Passive investment giant Vanguard is hiring data scientists to use analytics to provide insights into client needs. Similarly, BNY Mellon Investment Management's newly minted data-science team is looking to scour satellite images, web data and proprietary custody data to help its portfolio management, sales, marketing and strategy efforts. And BlackRock, the world's largest asset manager, established a new unit called the data science core in February. Among its aims will be to set "policy on algorithmic accountability and ethics of data science."

Like many asset managers pursuing a machine-driven approach to investing, it is Lawler's expectation that the future of investment management will always be a combination of man and machine. Few of those waving the flag of digital revolution think they are building the computers and algorithms that will one day replace them.

"We don't believe that all of this data science, which is quite powerful, is going to double people's Sharpe ratios, nor that there will be no room for humans. It takes a lot of human input to program the models and maintain them," he says. "However, what we do believe is that other people will be looking at data in very real time, so using data science is likely going to be necessary to deliver competitive performance."

And on this, he is resolute. "Whether you do it poorly or well," Lawler says, pausing, "that's data science." ■

Consolidated Tape Headed for a Regulation Delegation

Whether through an industry initiative or delegated acts, a consolidated tape provider will be established in Europe, and the industry could lose out if the regulators dictate the terms, reports Jamie Hyman.

Time is running out for an industry-led consolidated tape solution in Europe.

The revised Markets in Financial Instruments Directive (Mifid II) introduced the concept of a consolidated tape provider (CTP). According to Article 65 of the regulation, a CTP collects post-trade data published by trading venues and approved publication arrangements (APAs), and consolidates them into a continuous live data stream that it makes available to the public, both for equity and non-equity products.

The European Securities and Markets Authority (Esma) allows for up to two years following the Jan. 3, 2018, implementation of Mifid II for a CTP to step forward. If there are no takers, Esma is charged with reviewing the situation, and an Esma spokesperson confirms a review is scheduled for 2019, "but right now it is too early to be more specific on exact timing." If necessary, the regulator is empowered by Mifid II to mandate the establishment of one or more CTPs, but Esma declined to comment on whether that is an area currently under discussion.

"Delegated acts are on the horizon," says Graham Dick, head of client relationship management as Aquis Exchange, adding that ultimately, the establishment of a CTP may only be triggered by an Esma directive with input from various national competent

authorities (NCAs). "Esma will put something else out to tender, and once something goes out to tender, [the regulators] impose the conditions on which the consolidated tape is drawn up and done, rather than it being proposed by the industry."

Dick has a gimlet understanding of the difficulties inherent in establishing an industry-led CTP, because about five years ago, he was part of a partnership that spent a year and a half trying to create a consolidated tape business for pan-European equities.

"We gathered all of the vendors, all of the various providers of pan-European market data, the main incumbent exchanges, the main clients, and we spoke to a lot of the buy-side institutions to try and gather an industry consensus—and that had quite a lot of momentum in the early stages," he says. Although ultimately it did not work out, Dick says it is very clear that in order to have an "effective and efficient" consolidated tape, every participant has to be gathered and working together. "If one person gets away and doesn't want to participate, then your tape is less valuable. What happened is that we'd gathered quite a large number of providers but we hadn't gathered them all, and then some of the others started to flake away, and that's why it didn't work at the time."

The US Securities and Exchange Commission (SEC) established a consolidated tape in 1976, which reports price and volume data on exchange-listed stocks and could serve as a preview of what is to come in Europe. "When we see other markets that have a consolidated tape, like the US, it's nice to think we might be able to achieve something like that," says the head trader at a UK-based asset manager, who was quick to follow up that there are regulatory downsides to an electronic tape, such as the trade-through rule in the US, which requires orders to be executed at the lowest price on any market. "Sometimes the best price is not actually the best execution," the head trader says.

> "Getting that data becomes a very costly exercise, plus making sure that it's stable, reliable, putting all that together. With all of these fragmented sources of data, the cost of delivering that and producing it becomes too high." **Matthew Coupe, Barclays**

Data Matters

Certainly, the European market has done without a CTP for years, and so an argument could be made that the lack of a consolidated tape isn't really that big of a deal.

"If everyone is challenged, is anyone missing out?" asks Matthew Coupe, director of market structure at Barclays and co-chair of the EMEA regional committee and EMEA regulatory subcommittee for the FIX Trading Community.

The answer is unclear, and likely to change as the market evolves.

While the data may not be strictly necessary, it certainly could prove useful. The asset manager says if he could see what everyone else is doing in real time, he could determine "whether I'm a good trader or a bad trader because I'd be comparing my performance against a truer benchmark."

"Mifid is about making sure the end investor has a stable and transparent market to invest in, so we can actually get the right growth of assets. From a buy-side point of view, they want to be able to see and leverage that data," Coupe says.

Alex Wolcough
Appsbroker
Fintech

But perhaps the most valuable benefit of a CTP would be the potential for improved data quality.

One fintech consultant says APAs have an obligation to check the data and reject it if they find it is low-quality, but "I don't think they're doing an awful lot of that at the moment." He reports that data quality expectations were high for 2018, but when he looks at the raw data currently available, he spots multiple errors. He blames a current "regulatory imbalance" where regulators are tough on some APAs and trading venues but lenient on others. "I think that some of this is so bad that we're just stepping back to the world before Mifid," he says, adding that concern over data quality is high enough that end users are less willing to pay for the data.

Coupe says there are data quality issues, namely in post-trade data for the equities market and specifically when it comes to the identification of instruments within over-the-counter (OTC) derivatives. When looking at his organization's internal data, he adds, the key element is that it can't be materially misleading, and the APAs should be doing the same assessments.

"As this data is public, they should be making sure that the quality of the data they're delivering is accurate and provides a good understanding of what is actually happening, and

that they are identifying and flagging trades the right way," Coupe says.

The asset manager's quality concerns are focused on post-trade transparency. He says different data providers sometimes provide different volume-weighed opening price (VWOP) numbers, which may differ still from the number given by the third party providing his firm's transaction cost analysis (TCA). "It would be really useful to have a consistent number rather than everyone calculating things slightly differently," he says.

Coupe says FIX is setting up a committee to try to deliver technical standards that may solve some data inconsistency issues and smooth the path to a CTP.

"Hopefully after a bit of time we can get everyone writing to a single technology specification," he says. "What you find at the moment is one APA may have a different specification than another, generally because of the rapid nature of how things need to be delivered."

Coupe says there is no real market advantage to not having a technical specification or not enabling standardization of the data.

Money Pressure

Of course, it is impossible in the current industry landscape to talk about data without having a conversation about costs.

The optimistic view is that a CTP could mean lower data costs for firms, by driving down market data prices or allowing firms to cut back on outsourced services.

The pendulum swings both ways, however, and data costs are also seen as a major barrier to the establishment of a CTP.

"You're not just consolidating data, but you're consolidating costs as well," says Alex Wolcough, director of Appsbroker Fintech, whose team is "trying to aggregate as many different sources as possible for non-equity data to provide access to that market data in consolidated tape format." Like Dick, Wolcough worked on a project to create a consolidated tape several years ago while working for Reuters, which later became Thomson Reuters. He says a CTP would have to get data from all of the different suppliers, and they all want to charge for their product.

"Getting that data becomes a very costly exercise, plus making sure that it's stable, reliable, putting all that together," Coupe says, adding that to be viable, a CTP also needs to ensure that its delivery of the level of access Mifid II requires does not become a burden to users. "With all of these fragmented sources of data, the cost of delivering that and producing it becomes too high."

The major data providers appear to share the view that becoming a CTP does not make sense for their bottom lines. "We do not currently

intend to perform CTP activities, and will therefore not seek to register as a CTP," a Bloomberg spokesperson says, declining to elaborate.

A Thomson Reuters spokesperson also confirmed that the data giant has opted not to become a consolidated tape provider "at this time."

"As currently set out, we believe the CTP rules make it difficult for any provider to offer a commercially attractive CTP service that meets the needs of the marketplace," the Thomson Reuters spokesperson says, citing Mifid II obligations for the CTP to include data from all in-scope venues, including those with minimal liquidity, while the regulation fails to impose a corresponding obligation on the venues themselves.

Dick, however, sees a possible business case for a true CTP, developed as Mifid II intended. "The amount of data is huge, the industry has made progress but there's still quite a lot of dirty data, which means not everything is fully reflected properly, but it can be done," he says.

The first step, Dick says, is determining the licensing policies for each individual exchange's distribution of data. "I know it's post-trade, and I know it can be argued that 15-minute delayed data should be free, but there are entangled distribution licenses and all sorts of historical contracts between regulated exchanges and the big mega-data providers," he says, adding that once a CTP is mandated, Esma will have to demand that all data is made available to the CTP free of charge. "Then there is a nominal administration fee, which is low enough to be reasonable for everybody, but because it will have so much interest, a whole spectrum of people would be interested in subscribing, and I think you can make it economically viable at a low cost to the industry."

Coupe says he is concerned about a CTP passing expense along to data consumers.

Graham Dick
Aquis Exchange

"If there's an increased cost of ownership of the data and operating it, there's going to be an overall increase in the cost of management fees to the end firm, potentially," he says. "Is that the right outcome for the end investor? Probably not. People want to see this data if it's material to the instrument."

"I know Esma wants a consolidated tape," adds the asset manager. "I don't know how they would redistribute someone else's data without paying them for it."

Divide and Conquer

Breaking down the big project into those smaller market segments would be key, Dick says.

"If they want to address every asset class in Europe, to have some sort of massive consolidated tape, it's going to take years and years, and it's not going to be functional for a long time," he says. "I would say the equity market should be a relatively easy segment to start with, and then take the major countries and worry about the marginal 27th or 26th nation that has very little activity, if that's going to take longer to do. But get the main consolidated tape out, which represents 95 or 98 percent of all flows ... because if we start worrying about the marginal stuff around the edges, it's going to take forever."

Although he has a relatively optimistic views of the possibilities, Dick says he would have preferred an industry-led initiative.

"My humble opinion is that five years ago, the industry could have put this together and done it themselves. That would have been a consensus view, which would have been a way of taking this forward," he says. "When anything is driven by a regulator, it's likely to be clunky, slow, and difficult to use."

Coupe says he sees a couple of options moving forward. First, the regulators could deliver a CTP themselves, or with a technical partner. Second, Esma could change the regulation "and give the ability for the market data providers to stand up and become CTPs, and make it more viable for them to do so. But then you are giving greater control to the market data providers in terms of what feeds they should select and what data they shouldn't. It's a huge amount of control."

However it shakes out, Coupe says it is essential that the industry and regulators communicate while staying focused on Mifid II's intent to provide transparency to the market as a whole.

"Whether you're an APA, a market data provider, an investment bank, buy side or sell side, what we all want to see is a clear and accurate representation of what's happened in the marketplace. As soon as that data paints a different picture, we need to sit down as an industry, work through those challenges, and come to a solution. That also means working with regulators, which are themselves industry members. There are a number of conversations ongoing and that really helps to get a positive result," he says. "What we need is a consolidated tape that provides an effective result of what the market is trading, to the appropriate levels as described within the regulation, and which gets the data out there in the right way while also not creating too many barriers to entry, commercially. If you create that commercial barrier to entry, you're cutting out the small guy, who Mifid II is meant to be there for."

If and when Esma takes the wheel, the industry would certainly give up some control, but that might not be a bad thing. "It's always best to be self-regulating rather than dictated to, and it would be nice if the data providers could get their act together ahead of that, but they will, I'm sure, fight it because I think it will hurt them financially because they're able to charge so much," says the buy-side head of trading, who adds that delegated acts would probably be "a move in the right direction." ■

Human Capital

ISN Taps Industry Vets Iati, Bruno

San Francisco-based consultancy International Solutions Network has hired Bob Iati and Mike Bruno as managing directors, to expand its roster of senior industry executives. Iati spent the past six months as an independent consultant, prior to which he was senior director for capital markets at Dun & Bradstreet, and spent more than 10 years at research firm Tabb Group as partner and global head of consulting. Before that, he was research director at Tower Group, and held VP roles at Deutsche Bank Securities and Lehman Brothers.

Bruno was most recently SVP and head of product management for North America at Rimes Technologies, prior to which he worked as a consultant business analyst for Morgan Stanley Wealth Management and as a senior consultant in a prior stint at ISN. Before that, he spent five years at FTSE, including as director of fixed income, head of strategy for fixed income and alternatives, and fixed income business unit head, and spent almost nine years at Reuters America, including as VP of institutional fixed income, and VP of fixed income product specialists. He joined the vendor from Bridge Information Systems, where he was VP of client services, following its acquisition of fixed income pricing vendor EJV Partners, where he was an account manager.

Mike Bruno

Both are based in New York and report to senior managing director Michele Kelsey.

John Faulkner

Exegy Promotes Anderson, Hires Faulkner to Bolster Sales

St. Louis-based hardware ticker plant provider Exegy has promoted Robert Anderson to senior account manager and has hired John Faulkner as sales director for EMEA and Asia-Pacific, based in London.

Anderson, who is globally responsible for sales of the vendor's Exegy Hosted and Data Solutions services, has spent the past five years at Exegy in various roles, including director, product manager, business development and sales associate, and business analyst. During his tenure, he led Exegy's effort to become a vendor of record, and launched its new X-Port and Data Port tools. He reports to COO Rod Arbaugh.

Faulkner, who is responsible for sales of Exegy's products in EMEA and Asia, joins the vendor from low-latency infrastructure provider Fixnetix, where he spent four years, including as director of sales and account management, and as technical account manager, prior to which he was VP of BARX client services at Barclays Investment Bank. Before that, he was a client FIX connectivity specialist at Instinet, and spent four years at Bloomberg in various technical and trading support-related roles. At Exegy, he reports to Carlos Lopez Lansdowne, sales director for the EMEA and APAC regions.

Bob Iati

MDX Ordains Bishop for BizDev

Data technology provider MDX Technology recently hired Darren Bishop as a business development specialist in London, responsible for the vendor's Project Iowa ecosystem that connects content creators and consumers, and for driving business development and creating new product offerings.

Bishop was previously CEO of LiquidityChain, a joint venture between Formulate Digital and interdealer broker TP Icap, where—prior to the merger of Tullett Prebon and Icap—he spent four years at Tullett as director of Tullett Prebon Learning and head of customer relations for EMEA, and before that spent six years at Icap as founding partner of its ReMatch emerging markets credit default swap execution platform.

Fenergo Taps Clarke to Lead Alliances Team


Julian Clarke

Fenergo, a Dublin-based provider of client onboarding, lifecycle management, counterparty data management, anti-money laundering and Know Your Customer tools, recently hired Julian Clarke as global head of partners and alliances, responsible for leading a new team focused on developing the vendor's partner ecosystem. Clarke was most recently interim chief client officer at Infuse Consulting in London, and held a similar role at merged technology consultancies Certeco and P2 Consulting, prior to which he spent three years at Capgemini in roles including global director of digital assurance and testing. Before that, he was group service lines business development director at Groupe Steria, and was sales and marketing director at software quality testing company Experimentus.

In his new role, Clarke reports to Greg Watson, global head of sales at Fenergo.

He also served as partner at The Beast Apps for its corporate CDS execution business, and spent 18 years in a prior stint at Tullett Prebon as head of EMEA and Asia information sales, and at its predecessors Tullett and Tokyo and Tullett Liberty as a business manager in its interest rates division, and providing strategic support and development to its brokerage business.

At MDX, he reports to CEO Paul Watmough, who says Bishop's appointment will enable the vendor to exploit new opportunities.

CJC Taps Love, Moreton

IT managed service and product provider CJC has promoted senior technical director Steve Moreton to global head of product management, and has hired Karen Love as business development specialist.

Moreton's primary focus will be MosaicOA, a real-time infrastructure data processor.

Moreton has spent 14 years at CJC in various consultancy and internal roles, following a four-year stint at Thomson Reuters as installs engineer and specialist project manager. His consultancy roles include market data IT consultant at JP Morgan and UniCredit, market data support specialist at Commonwealth Bank of Australia, and RMDS market data consultant at Thomson Reuters. His internal CJC roles include director of operations for Asia and senior technical director of global accounts and product management.

As the firm's new business development specialist, Love will be responsible for client outreach and finding new areas for CJC software solutions. She has spent her career in a range of front-to-back-office positions, most recently as account director at Fixnetix, a technology provider of market data, low-latency trading, connectivity and hosting solutions. Her previous experience includes roles as director of sales and account management at trading technology and infrastructure provider Ullink, sales director at NYSE Euronext, global account director, sales director of global connectivity, and manager of broker relations at Fidessa, and vice president of emerging markets operations at Citigroup.


Karen Love

Both report to CJC CEO Paul Gow.

Vela Names Former NYSE , Bank Tech Exec Visconti CTO

New York-based data and trading technology vendor Vela has hired Scott Visconti as CTO, responsible for leading the vendor's global development team, and focusing on its technology strategy and supporting its growth to the next stage of the company's development.

Visconti joins Vela from data analytics and trade surveillance startup MarketOpen, where he was co-founder, prior to which he was SVP of IT at IntercontinentalExchange, and oversaw the integration of the NYSE Euronext technology organization with ICE. Before that, he was managing director of equities IT at BNP Paribas, spent seven and a half years as director of equity derivatives IT at Credit Suisse, and previously served as global system manager at S&P Comstock, and an IT software consultant at Bedford Associates.

Based in New York, Visconti reports to CEO Jennifer Nayar. He replaces former CTO Aaron Wald, who has left the company after overseeing the integration of the Object Trading and OptionsCity trading technology and market access businesses it acquired last year.

Safra Names Brazil Exchange Vet Castro CIO

Brazilian banking group Safra has hired Marcio Castro as CIO, respon-


Steve Moreton

sible for systems development across all its business functions, including brokerage, investment banking, treasury and asset management, as well as retail banking channels.

Castro was previously IT director at Brazilian exchange B3 (formerly BM&F Bovespa), prior to which he held CIO roles at Brazilian clearinghouse Cetip, XP Investments, and electronic payments platform Redecard, and was director of operations at Interactive Brokers in Brazil. Before that, he spent nine years at BM&F Bovespa in various roles, including CTO and IT officer for trading and market data systems. During this period, he also served as co-chair of the Latin America sub-committee of standards body FIX Protocol Ltd (now FIX Trading Community).

Before joining the exchange—where he also worked as a project manager between 1999 and 2000—he was a treasury systems consultant at BankBoston, was professional services director at SunGard Treasury Systems, head of strategic development at the short-lived Australian Derivatives Exchange, and spent six years at IBM as a senior services specialist.

Delphix Adds Former FirstRain CEO Herscher to Board

Cloud data management platform vendor Delphix has appointed fintech executive Penny Herscher to its board of directors, to help take the company to its next phase of growth.

Herscher, who holds a variety of board roles at technology companies, was president and CEO of artificial intelligence and business analytics platform vendor FirstRain from 2005 to 2015, and served as chairman until its 2017 sale to Ignite Technologies. Before that, Herscher was president and CEO of semiconductor design company Simplex, which she sold to Cadence Design Systems, where she then served as general manager and chief marketing officer. Previously, she was general manager and VP of marketing at Synopsys, worked in research and development at Daisy Systems, and began her career as a R&D engineer at Texas Instruments.

Anova Taps Net Vet Hilt for Sales, Marketing

Chicago-based wireless data network provider Anova Technologies has hired network industry veteran Joe Hilt as vice president of sales and marketing, responsible for supporting a global expansion into Europe and Asia and for rolling out new products for financial clients. Hilt was previously VP of sales for financial markets at GTT Communications, which he joined via the network provider's acquisition of long-haul undersea cable network operator Hibernia Networks, where he held a similar role and also served as VP of sales for the Americas, having originally joined Hibernia Metro as VP of sales and marketing in 2006. Before that, he was director of sales and marketing at KeySpan Communications, was director of wholesale business at WorldLink, and held account manager roles at IXnet and Worldcom.

Based in New York, Hilt reports to Anova CEO Michael Persico.


David Puth


Steve Moreton

CLS CEO Steps Down

David Puth, CEO of CLS, has stepped down after leading the settlement, processing and data solutions provider since 2012.

Chairman Kenneth Harvey will act as interim CEO, with Puth remaining on hand through November to ensure a smooth transition.

Puth led CLS through a number of transitions, including its designation of a systemically important financial market utility, the release of several new products and services, a rebranding and investments in CLS's technology infrastructure.

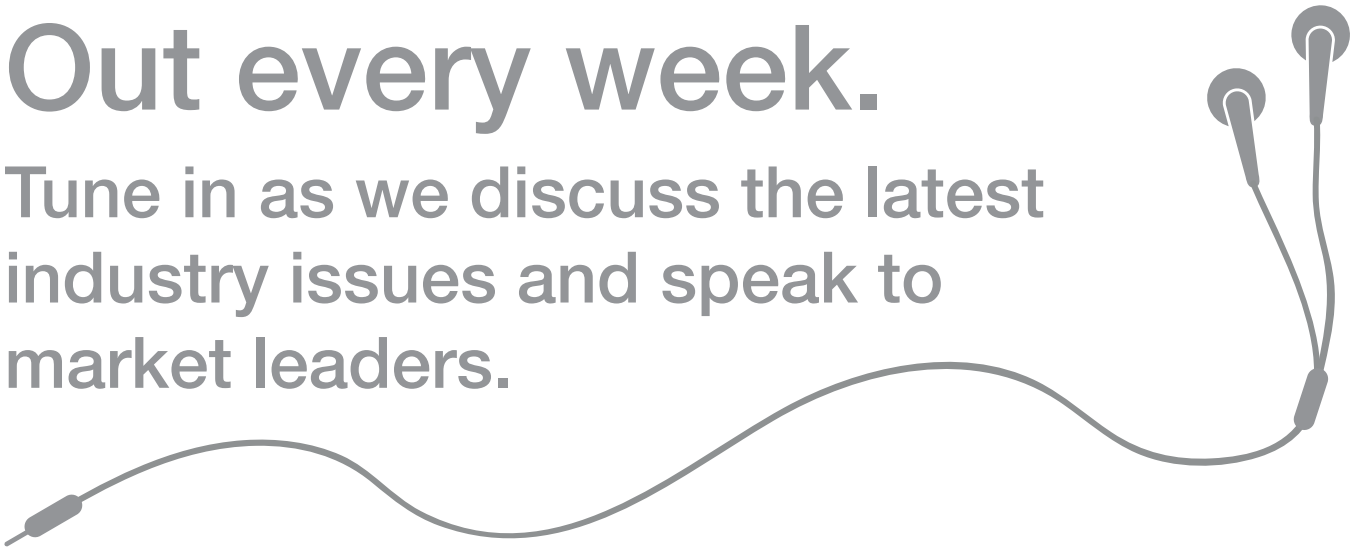
New York-based CLS has hired an executive search firm and is evaluating candidates for the CEO role. In the meantime, while Harvey is acting as interim CEO, CLS Independent Director Bryan Osmar and Director Rick Sears will assume the responsibilities of Chairmen of the Board of CLS Group Holdings AG and CLS Bank International.

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
or search the iTunes store for
"WatersTechnology"

RDU

The SmartStream Reference Data Utility

smartstreamrdu.com

Systematic Internaliser Registry. Centralized.

The Systematic Internaliser Registry (SI) provides the necessary data to determine which counterparty should report, by allowing SIs to register the details of the financial instruments and asset classes for which they are providing services, in a single centralized listing.

The SI Registry is a collaboration between the SmartStream RDU and a group of Approved Publication Arrangements (APAs). The initiative is open to all APAs and SIs and offers a comprehensive and granular set of data from a growing population of SIs - providing much needed clarity to the market.

Contact us today to find out more: **info@smartstreamrdu.com**