How Data WILL Determine the Future of RegTech

Donnelley Financial Solutions

Introduction

In heavily-regulated industries, such as financial services, health care, and insurance, "RegTech" is all the rage. Every major consulting firm has released a white paper on the subject. Over just the past two years, Google searches for the term have increased sevenfold.

What is RegTech, and why is it a newly-ascendant buzzword? The Institute of International Finance defines RegTech as "the use of new technologies to solve regulatory and compliance requirements more effectively and efficiently." Meanwhile, PwC says: "Rising compliance costs, along with regulators' and the industry's growing interest in automation, have created an environment ripe for disruption by emerging RegTechs, the innovative technologies that are addressing regulatory challenges in the financial services world."

RegTech solutions include new ways to 1) automate regulatory reporting, 2) derive insights from regulatory information, and 3) share information on complex markets and products. This paper describes four examples: in the first category, Donnelley Financial Solutions' ActiveDisclosure solution; in the second, idaciti and Intrinio; and in the third, TruSet.

All four of our examples use standardized data fields and formats to create new efficiencies. All four could deliver even greater efficiencies if U.S. regulatory agencies achieved better standardization. One of the key messages here is that *RegTech solutions require data standardization*.

Conversely, if regulatory agencies fail to make further progress in standardizing the data fields and formats for the information they collect from the industries that they regulate, RegTech solutions cannot deliver much by way of further growth.

Donnelley Financial Solutions Director of Business Development John Truzzolino explains that the future of RegTech, including Donnelley Financial Solutions ActiveDisclosure, rests on "the migration from [regulatory] disclosure of documents to disclosure using structured data." Craig Clay, president of global capital markets for Donnelley Financial Solutions, explains: "When RegTech connects with data... it's a way of creating transparency. It's a disruptive approach to solutions that rests on a few key themes: efficiency, minimizing risk, and improving quality."

Hudson Hollister, founder and executive director of the Data Coalition, a Washington, D.C.-based trade association, notes that today's regulatory environment is characterized by unprecedented complexity. As of late 2017, the Securities and Exchange Commission (SEC) alone used 600 separate forms to collect disclosures from public companies, financial firms, funds, and exchanges. The SEC has transformed a few of these disclosures from old-fashioned documents into standardized data—and those disclosures are the ones that now support RegTech solutions.

If the SEC transformed all 600 of its forms from documents into standardized data, and if other federal, state, and local regulatory agencies did the same—then a true transformation would be underway.

Hollister maintains that when data is missing from the RegTech equation, the power of the solutions is greatly diminished. Hollister emphasizes, "RegTech solutions require good, structured data. It's that simple."

Some of the innovators profiled in this paper developed their ideas before "RegTech" became a buzzword. Emily Huang, CEO and co-founder of idaciti, says that when she founded her company in 2014, "the term 'RegTech' wasn't even in the vocabulary … We never said: 'Hey, we want to play in the RegTech space.'" Huang developed a new way to use standardized public-company financial data to bring insights to investors, an idea that fits the phrase that only later achieved popularity.

Like Truzzolino and Hollister, Huang maintains that more attention must be paid to the standardization of regulatory data. When she co-founded idaciti, "the missing piece was not the availability of the [information], but rather the usability of the [information as] data. And we wanted to show how the data could play an important part" in delivering easier, better insights to investors.

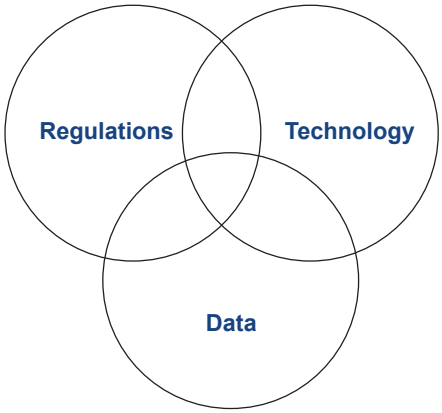
> *"When RegTech connects with data... it's a way of creating transparency. It's a disruptive approach to solutions that rests on a few key themes: efficiency, minimizing risk, and improving quality."* – Craig Clay, President, Global Capital Markets, Donnelley Financial Solutions

"RegTech solutions require good, structured data. It's that simple." – Hudson Hollister, Founder and Executive Director, Data Coalition

Key Takeaways

- RegTech solutions require data standardization.
- Regulatory agencies can maximize the promise of such solutions by coordinating changes in regulations, technology, and data.
- As structured data replaces old-fashioned, document-based disclosures, it ought to be subject to audit and quality requirements.
- In the long term, regulatory agencies can create an entirely new paradigm for RegTech by embracing a solution like Standard Business Reporting (SBR), which would serve as a common data structure for multiple regulators' regimes.

The Intersection of Regulations, Technology, and Data



If a Venn diagram were drawn with "regulations," "technology," and "data" as the three circles, the center is where new solutions scale fastest and deliver the greatest benefit. To deliver the most meaningful RegTech transformation, changes are needed in all three areas, and these changes must be coordinated with one another.

Standard Business Reporting, as practiced in Australia, the Netherlands, and elsewhere, shows what happens when government and industry work together to change regulations, technology, and data in a coordinated fashion.

Starting in 2008, the Australian government began adopting a single, standardized data structure for the information that its regulatory agencies collect from industry. In 2010, the government published its first version of a comprehensive taxonomy, or list of data fields, covering Australian companies' reports to multiple regulatory agencies. At the same time, the government worked with the Australian software industry to encourage tech companies to build software using these data fields to automate regulatory compliance. In addition, regulations were adjusted to align reporting procedures with a data-centric, rather than document-based, compliance model.

For Australia, these coordinated changes in regulations, technology, and data have created a true RegTech transformation. Australian software vendors used the standardized data structure to build new compliance solutions. Using these solutions, Australian companies can now comply with at least five different regulatory reporting regimes within one software environment. By the 2014-15 fiscal year, Standard Business Reporting was saving Australian companies over $1 billion per year through automation.

Australia's Standard Business Reporting program was patterned on an earlier, equally-ambitious program, also called Standard Business Reporting, in the Netherlands. The Dutch program also created a standardized data structure, coordinated with changes to regulations and the development of new technologies. In the Netherlands, the Standard Business Reporting regime is now the sole means by which many regulatory reports are submitted from industry to government. Meanwhile, the Estonian government has gone even further, adopting a common data structure for all government interactions by both companies and individual citizens.

The Australian and Dutch SBR programs both used the eXtensible Business Reporting Language (XBRL), a freely-available standard for exchanging business information. XBRL allows the expression of semantic meaning commonly required in business reporting. Each SBR program combined a common taxonomy of definitions used in regulatory reporting with XBRL to create a predictable electronic structure for compliance reports submitted by companies to government agencies.

For Australia, the Netherlands, and Estonia, the RegTech transformation offers more than just automatic compliance for businesses. It also dramatically improves the accuracy of information being reported because there are fewer opportunities for mistakes to be made during manual transcription. For example, once Dutch tax software vendors began using the Dutch taxonomy to automatically validate tax reports before submitting them to the Dutch Tax Administration, filers made more corrections before submission and the quality of final submissions to the agency improved.

Like its Australian and Dutch counterparts, and regulatory agencies in nearly thirty other countries around the world, the U.S. Securities and Exchange Commission uses XBRL to collect financial information. One meaningful, and long-anticipated, improvement for the SEC's regime is the adoption of the Inline XBRL, or iXBRL, format, for the disclosure forms that already involve some structured data. Inline XBRL is both human-readable and machine-readable, which means that a single document can be displayed on a browser for manual review and also ingested into software for electronic analysis. The SEC proposed a new rule to replace XBRL-based reporting with Inline XBRL in several of its disclosure forms in 2017. Final action is expected this year.

Even though Inline XBRL is much less ambitious than the multi-regulator SBR programs in other countries, it would be a good example of a well-coordinated change to regulations, technology, and data if it were formally adopted for some of the SEC's disclosure forms.

Until U.S. regulatory agencies get more ambitious, they will leave many potential RegTech solutions unexplored. Rachel Carpenter, co-founder and CEO of Intrinio, built her business with an intent to democratize financial market data access by cleaning up publicly-available regulatory data sets and making them available as easily-connected data feeds. But she says that Intrinio's ability to do this depends on "the regulatory environment increasing the amount of data sets that are filed digitally." Carpenter points out that while the SEC has mandated the use of XBRL data within some corporate disclosures, many other agencies have not adopted any data structure at all for the information they collect.

For example, if municipal governments and nonprofits—all of which generate government-mandated financial statements—made that data accessible in XBRL or even via an API, an Excel add-in, or Google Sheets, consumers could gain access to critical data in a much more flexible and open way, says Joey French, Carpenter's co-founder and President and COO of Intrinio.

French points out that hundreds of thousands of city financial reports are sitting in PDF files on tens of thousands of websites across the United States. "You can't get access to the data. Nobody can analyze a municipal bond. It's 2017 and we aren't filing that data digitally. It's insanity," says French.

> "*It's 2017 and we aren't filing [municipal bond] data digitally.*" − Joey French, Co-Founder, President and CFO of Intrinio

XBRL, Inline XBRL, and SBR

- **What is XBRL?** XBRL—eXtensible Business Reporting Language—is a freely available and global standard for exchanging business information. XBRL, a specification developed and published by XBRL International, Inc., allows the expression of semantic meaning for terms commonly required in business reporting.
- **What is Inline XBRL?** Inline XBRL, or iXBRL, allows structured data XBRL tags to be included behind the scenes in a human-readable format that is displayed on a browser instead of being located in a separate document. The structured data, which can easily be processed by analytical tools, is closely tied to the numbers and text presented within the human-readable format.
- **What is SBR?** Standard Business Reporting, or SBR, starts with XBRL and other syntaxes and incorporates a taxonomy of definitions used in government legislation and reporting; these harmonized terms are then linked to the same exact standardized terms within business and local accounting software. The history of SBR began with the Netherland's Taxonomy Project in 2004, and in 2008, Australia and the Netherlands formed the SBR International Forum.

Even when regulations, technology, and data are changed in a coordinated way, data quality remains crucial.

When it comes to financial statement data, global financial markets trust the information provided because such statements are audited. According to the CFA Institute, investors are surprised when they learn that financial statements delivered digitally are not audited. Indeed, 50 percent of CFA Institute members surveyed in 2016 believe that digital information should be incorporated into the standard financial statement audit. Accordingly, Truzzolino asserts that a global move to audit iXBRL is necessary in order to ensure the trustworthiness of digital financial data. He notes that accurate, audited digital financial statements would facilitate analysis and could minimize errors in the translation from HTML to digital versions, enhancing the usability and quality of the digital data being collected.

Until recently, assumptions about the way users consumed financial statement data seemed beyond question: Users read documents, end of story. In reality, the majority of financial statement data today is consumed digitally.

When asked why financial data should be digitized, the answer is "usefulness." In a machine-readable and –consumable format, financial data can be used for far more sophisticated analysis on a company-specific basis—or across industries or even the entire population of companies out there. Doing this analysis, however, rests on data being *reliably* digitized according to approved data definitions and data standards.

Prominent voices in the industry have also argued the case for data quality. According to idaciti's Huang, "This is something I say all the time: Just because data is available doesn't mean data is usable.'"

> "*This is something I say all the time: 'Just because data is available doesn't mean data is usable.'*" – Emily Huang, CEO and Co-Founder of idaciti, Inc.

Case Studies: Four RegTech Solutions

As the last section suggests, RegTech solutions succeed when they apply technology to process data to deliver new efficiencies, in a manner consistent with regulations. Such new efficiencies can include automated regulatory reporting, faster or better insights from regulatory information, and/or shared (and therefore cheaper) information flows for complex markets and products.

The regulatory push for data quality has come as a clarion call, but specific innovations have arisen from RegTech companies themselves, including the four profiled in the following case studies.

From Donnelley Financial Solutions' (DFS's) ActiveDisclosure, which allows filers to populate regulatory submissions automatically, to idaciti, Inc., a software solution provider that offers normalized structured financial data and auto-tags additional non-financial data within financial disclosures, these emerging solutions are furthering a coordinated RegTech Data vision. Intrinio, creator of a RegTech Data marketplace, is helping disseminate machine-readable information to a broader audience, while TruSet is testing the limits of how blockchain can help a community of users contribute to a common set of machine-readable data for fixed-income prospectuses.

For now, each of these solutions takes advantage of regulatory data currently available, applying its own technological innovations, and complying with current regulations.

Donnelley Financial Solutions' ActiveLink application, for instance, pulls reporting data directly from Excel into an SEC disclosure document, promoting accuracy by eliminating error-prone reformatting, cutting, and pasting. "Clients have confidence that the numbers in the SEC filing came right from their financial reports," says Darren Peterson, DFS's Senior Software Product Leader.

Will Janensch, Co-Founder and COO of TruSet, is using blockchain technology to improve data quality by inviting a community of users to come together to "cleanse" the structured data around fixed-income instruments. "One of the keys to making any kind of automation work is that the inputs that go into the system need to be correct and trusted," says Janensch.

But if the SEC and other U.S. regulatory agencies further modernize their regulations and transform more document-based disclosures into data, while at the same time ensuring the quality of that data, many new possibilities will open up.

If regulatory agencies replace more of their disclosure documents with structured data, then "[t]he information reported to the

agencies can be used by the agencies immediately, without having to correct or change or question the data being collected," says DFS's Truzzolino. At the same time, he explains, users of RegTech solutions like ActiveDisclosure will shave hours and days off validating data and can instead "spend more time on the management of the company and the analysis of the data that's been collected."

> "*One of the keys to making any kind of automation work is that the inputs that go into the system need to be correct and trusted.*" – Will Janensch, Co-Founder and COO of TruSet.

1. Automated Reporting: ActiveDisclosure

What if a filing solution for corporate issuers that relies on good structured data could also help issuers generate higher quality data that could be validated in real time?

Since the first half of 2013, when ActiveDisclosure, Donnelley Financial Solutions' disclosure management SaaS application, was publicly released, issuers mandated by the SEC to create machine-readable data can meet their regulatory commitments more simply and more accurately, according to Peterson.

ActiveDisclosure is able to provide more than just compliance documents because the SEC decided to adopt XBRL as its standardized data format for financial statements. Should the SEC adopt standardized data for the remainder of its required corporate disclosures, then ActiveDisclosure would be able to automate an array of additional tasks that today require manual compliance.

Issuers prepare XBRL filings within ActiveDisclosure, which is a collaboration platform for finance and SEC reporting teams who prepare quarterly and annual reports. The solution ensures that important steps along the filing journey are successfully completed, explains Peterson.

It's also increasingly clear that automated reporting software can help resolve some of the problems within the RegTech Data paradigm. Take, for instance, the nagging problem of data quality. Peterson points out that ActiveDisclosure software resembles the concept of straight-through processing embraced by other industries because it is built on the notion that data should flow from the source system to the destination without the need to manipulate it by hand. He continues: "As soon as you [manually] touch data, whether it's in Excel or a desktop, there's the opportunity for errors to be introduced."

A push for data quality has direct implications for workplace efficiency, as ActiveDisclosure users can attest. "We've had clients comment that they spend as many as 800 person-hours per year 'ticking and tying' numbers to ensure they are accurate back to the source locations/systems," says Peterson. "That's one of the key reasons why ActiveDisclosure was designed to link to those sources rather than attempt to replace or replicate them."

Going forward, DFS seeks ways to integrate its existing solution so that it can help fulfill other compliance needs, as well. Peterson emphasizes that much of compliance revolves around Sarbanes-Oxley Act compliance, so DFS is now including AuditBoard's SOXHub technology, with all its references and checklists, into the process of creating SEC reports.

Ranging from private companies to government entities, any organization that produces large documents containing data and narrative content can benefit from the ActiveDisclosure platform, says Peterson. That's because ActiveDisclosure brings structure and visibility to large, multiple-contributor content projects.

2. Faster, Better Insights: idaciti and Intrinio

What if machine-readable data could be made available for parts of financial documents that are typically not tagged, allowing business insights to flourish and meaningful comparisons to be made between different companies and industries?

For Huang, the key is looking beyond the data that is currently being tagged and making all existing data available for a much wider array of purposes. By building on the SEC's existing foundation of requiring that structured data be supplied for financial statements, idaciti, a software tool that facilitates the accessing, analyzing, and visualizing of financial and non-financial data, has become a platform that can help users make sense of unstructured content, as well. Thus, idaciti uses technological innovation to extend the bounds of what is possible with currently-available, machine-readable data.

"When I read a 10-K, I go right to the MD&A [Management Discussion & Analysis]. That's where management discusses how they performed and answers the important 'why' for their performance," says Huang, noting that the SEC does not require the MD&A to be tagged. "So if you look at what XBRL has covered, it's a lot of 'what.' But the answer to 'how' and 'why' are in the unstructured portions of the documents."

Idaciti has created a software solution that uses machine learning to auto-tag unstructured portions of financial filings and capture the valuable insights in parts of a financial submission that were previously overlooked in the tagging process.

This solution also extends the existing XBRL taxonomy to key performance indicators (KPIs) that may matter within a particular industry but are not captured by the XBRL tagging process. Huang points out that revenue, which is part of the XBRL taxonomy, is not necessarily a KPI for a social media company, while the number of active users and the number of active users on mobile devices *are* valuable pieces of information (and yet the SEC has not required these metrics be tagged). When idaciti's software captures these KPIs in a machine-readable format, then meaningful comparisons can begin to be made among social media companies.

XBRL-tagged information can be invaluable when there are new accounting standards such as those around revenue recognition. While companies like Microsoft pioneered the adoption of new revenue recognition standards—and even discussed the process for doing so in SEC filings—most other companies did not reap the benefits of Microsoft's early adoption experience because the data was buried.

"When searching through all public company filings, it is impossible to locate a concept like who early-adopted revenue recognition," says Huang. "Yet companies also want to know: What kind of money and time did a company spend to adopt revenue recognition? And overall what was the impact of adoption?" Once sections of SEC filings on revenue recognition and other processes are tagged, the information will be readily available to all interested parties.

"For us, it's not just about the numbers provided to the regulator," says Huang. "A lot of narratives included in the filing truly can help the company learn from what other companies are doing and what are the best practices."

As an example of how idaciti might make a difference, Huang points out that an oil and gas company that historically spent 640 hours a year comparing its data with 25 other companies found that those comparisons could be made in a matter of minutes with idaciti. More importantly, though, the idaciti platform led to a shift in how this company benchmarked itself against peers beyond the original group of 25. Huang suggests this shift is critical: "This oil and gas company is now thinking about what possibly can be done, rather than just focusing on what it's able to do."

Huang is convinced that regulators need solutions like idaciti's platforms as a way of showing that tagging data in XBRL can have enormous strategic pay-offs. If the regulators could show the potential benefits from comparing a company to its peers or how companies are handling the adoption of new accounting standards, then they could spotlight the true value in digitizing financial data.

"With RegTech, unless you can demonstrate the benefit and quantify the effectiveness a company can get from the technology, you can't move forward," says Huang. "Knowing that 640 hours can be reduced to minutes—or the impossible task of finding the early adopters of an important accounting policy standard can be accomplished—is what needs to be showcased."

Intrinio, which has built a marketplace for over 200 data sets, including XBRL-based ones, also automates insights by allowing users to connect usable, quality data in groundbreaking ways. This RegTech solution exemplifies the value of high-quality, structured data by cleaning up existing data sets from the SEC and other sources and making them far more widely available to users.

Carpenter points out that roughly half of Intrinio's users are investors, eager to analyze financial statements, and half are developers seeking access to quality data in order to innovate by developing mobile and web apps and risk analytics software, among other things.

She describes Intrinio, which has a marketplace that has grown to 20,000 users within just the past two years, as "the Amazon of financial data." Carpenter continues, "It's a website that you can go to and shop around for whatever data you need. You have a user profile and a shopping page, where you can basically scroll through all the different types of data available and access them."

As an example of the type of innovation that Intrinio is facilitating, Carpenter cites academic research. At a quant modeling class at the University of Tampa, students were attempting to analyze the price-to-earnings ratios for 500 tickers. Doing this manually meant going to the Internet and copying p/e ratios into Excel spreadsheets for all 500 tickers. Once students had access to an Excel product available on Intrinio, they simply dragged down a formula, populating all of the data automatically. "Because of XBRL and because of these mandates, we were able to source and clean up the data and get it into the hands of students, saving them a lot of time," emphasizes Carpenter.

As another example, French cites an Alexa application designed to answer financial questions ranging from the address of Amazon's corporate headquarters to the p/e ratio for Apple. He points out that when such an application is fueled by data from a marketplace like Intrinio's, a user can basically ask Alexa any question about a publicly-traded company and receive an answer in seconds.

French believes that good, structured data must be widely available. "The Holy Grail for us," concludes French, "is that the more types of data are filed digitally, the better—at the federal level, at the municipal level, and across different asset classes."

As valuable as solutions like idaciti's and Intrinio's have become, their value would only be enhanced if the SEC and other regulators chose to continue the transformation of regulatory reports from documents into data.

"*An oil and gas company that typically spent 640 hours a year comparing its data with 25 other companies found that those comparisons could be made in a matter of minutes with idaciti.*" – Emily Huang, Co-Founder and CEO of idaciti

"*The Holy Grail for us is that the more types of data are filed digitally, the better – at the federal level, at the municipal level, and across different asset classes.*" – Joey French, Co-Founder, President and CFO of Intrinio

3. Shared Information Flows: TruSet

What if there were a way for data users to correct mistakes communally, allowing each participant in a market to spot a problem and then rectify it for all other users of the data, too?

TruSet, a start-up seeking to help bond investors share intelligence from prospectuses, is using blockchain technology to achieve that vision. "Our solution," says Janensch, "will allow you to have more accurate data that you can better trust and save money in the process."

Janensch explains that TruSet addresses what is essentially an outdated regulation: the SEC requires all issuers of fixed-income instruments—or bonds—to generate and file in HTML lengthy prospectuses that are then published on the EDGAR site. However, the institutions reading these prospectuses and investing in these bonds need the information to be presented in a structured data format, so that they can plug information into their software and run all the necessary analytics.

Enter "the middleman," or large information vendors, such as Bloomberg and Thomson Reuters, that convert fixed-income prospectuses into a machine-readable format. Janensch explains that the middlemen rely on "brute strength," automatically scraping PDFs and employing teams of hundreds of individuals who help interpret the prospectuses.

This process is expensive, and "there are enough errors that the customers of those data feeds don't trust the data as accurately representing what those prospectuses said," explains Janensch. Because of these shortcomings, he finds that each asset manager uses software analytics and back-office personnel to "interrogate the data feed" in order to locate and correct errors, creating a so-called "golden record" that they deem trustworthy. Janensch characterizes the result as "a very inefficient market ecosystem."

Fortunately, blockchain technology has the potential to up-end the paradigm. Using blockchain, TruSet has built a prototype for a solution in which a community of users corrects data that is then shared among participants.

Janensch explains that a blockchain has three important characteristics: 1) it's a distributed ledger or database; 2) there is consensus around changes made to the database; and 3) there is cryptographic immutability.

Here is how each of these blockchain features drives the TruSet solution. Although not necessarily a true distributive ledger application in the tradition of Bitcoin, TruSet's solution is "a database that's distributed among all of the different consumers of the data," says Janensch. Specifically, the community of users delves into the database, identifying errors in any reference data and only accepting the final database information that is deemed correct. Instead of participants each having to cleanse the same exact data, they work together to create a final product that they trust is accurate.

Second, consensus for TruSet takes the form of a protocol that governs under what circumstances a change to the database can occur—and whether the community can reach agreement about the accuracy of that change. "We're using that consensus process to basically crowd-source what's now currently happening in silos, which is the data correction piece happening at the customer end," says Janensch.

The third important feature is cryptographic immutability. Although the distributed database can be changed, whatever happened is still recorded within the chain. In other words, within the blockchain exists a snapshot of all of the information in the database at any given instant.

While TruSet has built its solution around the Ethereum blockhain, it is working with a permissioned (or private) version that only allows participation by pre-cleared entities. These entities become nodes, and in this peer-to-peer community, a node may play one of three roles: publisher, validator, or consumer.

One of the ironies of the fixed-income world today is that the SEC requires financial institutions to produce prospectuses as unstructured documents; however, the agency also needs machine-readable data, so it buys that data from the current large data vendors. Were the SEC to collect fixed-income prospectuses as machine-readable data in the first place, this problem would disappear.

Prototypes like TruSet are filling a gaping hole by using technology to do what the regulators are not currently doing themselves. "One of the cool things about blockchain is it allows regulators to view and maybe act on things in real time instead of waiting until after the fact to report," says Janensch.

While TruSet is debuting a fixed-income solution, the same type of blockchain solution could be created for other instruments and industries that need shared reference data and are plagued by consistency and accuracy problems.

Janensch sees enormous potential for the TruSet model because data created by financial institutions on the TruSet platform would no longer be owned by large information vendors. He anticipates a time in which a bank issuing a bond might benefit from the data it generates by getting paid a fee for contributing records. He also foresees that those validators creating golden records might no longer *pay* to have the data cleansed, but instead might *be paid* for contributing value to the overall database.

Because data vendors "own" and charge for the data they render machine-readable, vendors both charge their customers for each business function for which they use the data and impose strict usage limits on that data. Janensch hopes that this, too, might change.

The current model, he says, "has retarded innovation around the data by not allowing the community that generated the data to create more sophisticated data services." In the TruSet vision, Janensch sees fixed-income data becoming "a community-owned resource" with far fewer restrictions on how the data is used and what asset managers can do with the data that they are actively working to improve.

"*One of the cool things about blockchain is it allows regulators to view and maybe act on things in real time instead of waiting until after the fact to report,*" says Will Janensch, Co-Founder and COO of TruSet

WHAT THE FUTURE HOLDS

In the United States, RegTech solutions will expand to the extent that the SEC and other regulators carefully coordinate changes in regulations, technology, and data. And as data replaces documents, it must be fully audited and of reliable quality.

1. The Need for Proper Auditing

In October 2017, the SEC approved a new PCAOB rule that requires significant enhancements to public company audit reports, including the disclosure of auditor tenure. These enhanced requirements became effective for audits of financial statements for fiscal years ending on or after December 15, 2017.

Arguably, the most significant change to the auditor's report is the communication of critical audit matters (CAMs), which will be mandated beginning on June 30, 2019. CAMs are matters that have been communicated to the audit committee, are related to accounts or disclosures that are material to the financial statements, and involve especially challenging, subjective, or complex auditor judgment.

"The changes adopted today breathe life into the audit report and give investors the information they've been asking for from auditors," said PCAOB Chairman James R. Doty.

Although the push for greater auditor accountability is welcome, regulators have not extended the new requirements to data submitted digitally. DFS's Truzzolino believes that overlooking structured data in this regard is a serious oversight.

Truzzolino is convinced that iXBRL could play a central role in the depth and relevance of information that users of structured data can access. A tag in iXBRL can, for instance, include a wealth of information about the data disclosed, and there is no reason why a tag should not also contain an auditor imprimatur, a link to audit guidance, and references to reported CAMs.

2. An Emphasis on Quality

Today, the financial data that companies and agencies routinely access to close their books is collected in a matter of days, and yet these companies and agencies take weeks to publish reports, delaying management and stakeholder analyses and decisions.

One reason for such delays is that information is contained in data warehouses or consolidation applications, where data is commonly cut and pasted, re-keyed, or manually transferred into word processing and spreadsheet applications.

Data standardization and the effective implementation of disclosure management applications can enhance and streamline this entire reporting process. Truzzolino notes that disclosure management applications provide report-writer functionality through word processing and spreadsheet applications commonly used in manual reporting steps. When data is standardized in this way, applications are able to pull information from disparate data sources to create automated reports.

As SEC Commissioner Kara Stein has stated, "improving the quality of data available on smaller and medium size companies could lead to improved secondary market liquidity. Improved data and transparency on market quality statistics could empower small and large investors and benefit the market overall. In short, the digital revolution is requiring us to rethink and re-envision disclosure."

In the end, RegTech Data—with an emphasis on "data"—has the potential to combine regulation/policy making, disruptive technology and data standards to streamline financial reporting, while enhancing data quality and making this digital financial data more usable for all stakeholders.

3. A Roadmap for Regulations, Technology, and Data

Following are steps that remain to be taken to achieve the RegTech Data promise in three critical realms:

Regulations. Truzzolino notes that custom extensions to the SEC's US GAAP taxonomy created by individual companies make it difficult for meaningful data comparisons between the various companies out there. He says that until there is true standardization within the structured reporting, achieving the original SEC vision "of leveling the playing field between companies large and small" for how information is presented and consumed by investment analysts will remain an elusive goal.

What's more, the SEC needs to fix the lack of comparability across its current XBRL-formatted financial statement submissions, contends Hollister. He is eager for "the SEC and

other regulators [to] adopt a standardized data structure for all of the information they collect from the financial industry."

Technology. Intrinio's Carpenter and French point out that software developers attempting to devise new solutions need quality data in a machine-readable format in order to build the solutions of the future. Without this data, innovation will be stifled.

Huang maintains that when new applications are introduced, it is critical that the practical uses be highlighted. As a cautionary tale, she notes that when the SEC unveiled its inline viewer for iXBRL, the regulator showcased the metadata, what she calls "the really geeky stuff," rather than the functionalities that could truly help investors.

The inline XBRL viewer enhanced by idaciti allows a user to click on a revenue number for a given company and then see not only the revenue reporting for that year—but a trending chart depicting whether revenue is increasing or decreasing. Huang notes that the inline viewer also makes it easy to benchmark any reported item for several companies at once. "The inline viewer brings the data to life, and people can see that this is the power of XBRL," she says. "The SEC has created a great foundation, but it's up to the agile software companies in the marketplace to add additional, innovative functionalities."

Huang is adamant that all technology should be presented in a way that makes the value to users clear. "It's important to show how we maximize the value and effectiveness for the issuer by automating using machine learning and other advanced technologies," she says. "We want to show how you can ask a question and get an answer without taking 17 steps to get there."

Data. As Hollister points out, many RegTech Data applications "are held back by a lack of accurate data." Making machine-readable data dependable is essential for true progress to be made.

When Intrinio's Carpenter envisions the RegTech Data future, she says that the goal should be "a superhighway" in which reliable and accurate data is generated by companies and flows instantaneously to users and innovators alike. With this data, companies, investors, and software developers will have the infrastructure necessary to gain needed insights for breakthrough developments.

In the short term, this means the SEC should adopt the Inline XBRL format for the financial statements that it already collects in XBRL. While in March 2017 the SEC signaled that it will mandate iXBRL as the reporting language for financial statements, that requirement has not yet taken effect. The SEC continues to require an HTML copy (in a human-readable format) and a separate exhibit in XBRL (as the machine-readable format).

Once an iXBRL mandate occurs for public companies in the United States, then those companies will no longer have to submit to the SEC the HTML and XBRL versions of the exact same filing. This shift will take some of the burden off issuers themselves, and it will result in greater accuracy because there will be a single filing rather than two formats that could have discrepancies between them.

In the medium term, the SEC and other regulators need to replace existing disclosures with standardized and structured data, encourage the development of both public-sector and private-sector technology to take advantage of that data, and adjust regulations to permit all manner of compliance to take place in a data-centric, rather than document-based, manner.

In the long term, many industry experts are eager for regulators to adopt an even broader and more far-reaching standardized data structure, such as Standard Business Reporting. Hollister points out that SBR would allow data collected from all corners of the government to be used in meaningful comparisons, spurring dramatic new RegTech applications.

> "*We want to show how you can ask a question and get an answer without 17 steps in the middle.*" – Emily Huang, Co-Founder and CEO of idaciti.

4. A Future Defined by Data

Hollister and Truzzolino point out that with the Financial Transparency Act and other Congressional developments afoot, now is a propitious moment to take a closer look at how data collection can be standardized across government agencies.

Meanwhile, Hollister notes that the Data Coalition is seeking reforms in Congress that would require financial regulators to adopt standardized data formats across all the information that they collect. Specifically, he hopes to see SBR adopted in the United States as it has been in the Netherlands and Australia.

"SBR," says Hollister, "is a standardized data structure that the United States should adopt for all the information that companies report to regulatory agencies." He continues: "Writ large, SBR is the culmination of RegTech Data. It's not possible to scale RegTech Data solutions unless you have data standards, and that's why we are pursuing the adoption of data standards by government regulatory agencies."

The goals of RegTech Data and SBR are sweeping but increasingly within reach. And while SBR looks promising, it is not the specific solution that matters so much as the overall commitment by regulators to demanding structured data that is presented in a consistent and easy-to-use way. The move from documents to structured data appears inevitable, but regulators will continue to play a critical role in helping achieve the full potential of the RegTech Data vision.

In the coming months and years, what will almost certainly be central to the success of RegTech Data is ensuring that "data" becomes an integral part of all facets of the RegTech conversation.

"Talking about 'Reg' and 'Tech' isn't enough," concludes Hollister. "Data is key."

WP-RegTechData 02.2018

Corporate Headquarters
35 West Wacker Drive
Chicago, IL 60601

800.424.9001

Donnelley Financial Solutions