



Inside Reference Data

October 2015

waterstechnology.com/ird

Data Quality

Special Report



Sponsored by

EAGLE
INVESTMENT SYSTEMS



Financial Information

Global Data. Swiss Quality.

Since 1930.

SIX Financial Information has prospered and innovated for more than 85 years by delivering the wealth and asset management industry with the highest-quality data and services. Our financial ecosystem is comprised of 16 million instruments from over 1500 worldwide sources, supplying financial decision makers with unmatched access to reference data, market data, corporate actions, regulatory data, and value-added services.

Elevate your client satisfaction with us.
www.six-financial-information.com



Financial Information

Unlocking the potential.

Editor's Letter



Control and Automation

For this report, *Inside Reference Data* asked participants in the Virtual Roundtable feature (page 8) how data sourcing issues can help or hinder data quality. Their responses all seemed to drill down to the issue of ownership of the data.

BNY Mellon's Amy Harkins says the biggest challenge when sourcing data in a large firm is putting its assessment in the hands of a single business owner. Eagle Investment Systems' Marc Rubinfeld identifies control of the sourcing as the greater issue. Control of the data, he says, makes it possible to enrich, repurpose or otherwise apply it to reporting, benchmarking or analysis. Firms should "have complete control of it and truly own it," he says.

Looking at another issue that can affect data quality—the automation of data processing—it seems it doesn't necessarily mean that quality will be decreased. Rubinfeld sees automation as "critically important to data quality, as it helps to discover errors that wouldn't otherwise be discernible," he says.

HSBC's Chris Johnson identifies another benefit of automation—getting faster data validation checks. SIX Financial Information's Dominique Tanner says when there are deficiencies in automated data processing, they can be corrected so they do not reappear. He brings the question around to sourcing, noting that firms must understand how sources are delivering data to properly map data to the correct fields. That's the precursor to automation.

A data-centric processing and delivery model or a centralized data team could be the best ways to raise the quality of the data being produced. That appears to be the frame under which ownership and automation should be sorted out.

Yours sincerely,

A handwritten signature in black ink that reads "Michael Shashoua". The signature is fluid and cursive.

Michael Shashoua
Editor, *Inside Reference Data*
Email: michael.shashoua@incisivemedia.com
Tel: +1 646 490 3969

1

IRD APP

THE INSIDE REFERENCE DATA APP GREAT NEW FEATURES

DOWNLOAD OUR LATEST UPDATE TODAY TO:

- **Get the full picture on the move** – text and images now automatically adjust to fit all screen sizes perfectly
- **Access the app on your desktop** – our new **Web app** allows you to view it on your computer as well as your smartphone and tablet



FEATURES

8 Virtual Roundtable

Inside Reference Data gathers together leading data management professionals to discuss how initiatives in sourcing, automation and processing can help improve data quality

18 Q&A

Data quality cannot be addressed with a broad governance sweep and should be seen from consumers' point of view, says Roberto Maranca, managing director for enterprise data at GE Capital

NEWS

6 Data Quality Concerns Rise on the Buy Side

6 Data Quality Efforts Focus on Sourcing, Efficiency

7 Consumers, Providers Stress Continued Importance of Data Quality

7 News Download



© 2015 Incisive Risk Information (IP) Ltd. Unauthorized photocopying or facsimile distribution of this copyrighted newsletter is prohibited. All rights reserved. ISSN: 1750-8517



Incisive Media
28–29 Haymarket,
London SW1Y 4RX

Data Quality Concerns Rise On Buy Side

Data quality is a top priority for the buy side—and it is rapidly gaining in importance, a recent survey has found.

Rimes Technologies recently released its third annual survey of market participants such as major asset managers, pension funds, banks, hedge funds and insurance companies. It canvassed 121 respondents in 41 financial centers across the world on their approaches to data governance.

Sixty-seven percent of respondents—24 percent more than in 2014—cited data quality improvement as the

primary data management priority for this year, says Giles Arbuthnott, Benchmark Data Service manager at Rimes. “This confirmed what we have been hearing in forums we held last year, and in industry research we’ve been reading: that firms now realize data quality isn’t necessarily the same thing for the end-users as it is for the operations teams working under more centralized models of data governance,” he says. “End-users are demanding better quality data that is fit for purpose.

Joanna Wright

Data Quality Efforts Focus on Sourcing, Efficiency

Sourcing and efficiency are proving big factors in data quality efforts by industry firms and their service providers, said panelists in a recent webcast co-sponsored by SIX Financial Information.

Having “exactly one source” for data to be managed is the most important piece of Goldman Sachs’ data strategy, said Gururaj Krishnan, vice president of futures execution technology at the firm. “Downstream consumers really reflect what the central source provides,” he said. “In many cases, downstream systems are doing their own enrichment, adaptation and transformation.”

Coordinating sources and operations

units is necessary to deliver data quality for the business, explained Dennis Gonzalez, head of platform re-engineering at consultancy Alta Strategic. “Without an initial buy-in, you end up starting a project and then often you don’t have the support you really need, and the same vision and goals that everyone can rally around,” he said.

Regulation, operational efficiency and business growth are the main drivers SIX clients cite when asked about data quality, added Dominique Tanner, head of business development at SIX Financial Information.

Michael Shashoua

Consumers, Providers Stress Continued Importance of Data Quality

Data quality should remain important despite a greater emphasis on cost savings and efficiency in the industry, say data users and providers.

“With the work that needs to be done with old data, old processes and the goals of our clients, we shouldn’t sweep all that under the rug,” says Adrienne Stern, region head of data change management for the Americas at SIX Financial Information.

Regulation will remain a driver of data quality, says Michael Niras, manager of the data management group at CIBC. “Firms collect data for two reasons—either to meet regulatory reporting requirements or to maintain a driver,” he says. “Where five big banks are saying you need to have your data this way, it doesn’t matter, because if you’re not in one of these two buckets of either maintaining and driving revenue for an institution or pushing on regulatory requirements and reporting, you will fall out of favor with your investors, shareholders and C-level executives.”

To achieve higher-quality data, the first step is to define what you want, states Adrian Dokmecian, data quality lead for hub and enterprise architecture at BMO Financial Group. “We basically want to understand what data is applicable to our users,” he says.

Michael Shashoua

Datactics Adds Data Integration Capability

Software vendor Datactics has launched a browser-based application called Data Quality Manager (DQM) for matching and integrating entity, instrument and regulatory data. DQM automates the transfer of critical data in the business life cycle and allows easier access to the data-processing engines of Datactics’ FlowDesigner application. DQM is compatible with multiple data formats and source information in the areas of entity, compliance and instrument data.

DTCC Repository Service Adds Testing Capability

The Depository Trust & Clearing Corporation (DTCC) has partnered with Risk Focus, a custom risk management solutions provider, to allow users of Risk Focus’s Validate. Trade validation and emulation engine to test-trade submissions in the DTCC’s Global Trade Repository (GTR) service, says Brian Lynch, CEO of Risk Focus.

Under the partnership arrangement, GTR users will get feedback on how trade messages should appear before being posted live on the GTR service.

Quality Controllers

Inside Reference Data gathers together leading data management professionals to discuss how initiatives in sourcing, automation and processing can help improve data quality

What is the biggest issue in sourcing data? Is it reconciling data from multiple sources, having the resources to get all possible sources or something else?

Amy Harkins, senior vice president and managing director, and head of enterprise client onboarding and global tax operations, BNY Mellon: The biggest challenge for sourcing the data in a larger organization is ensuring there is a single business owner to complete the needs assessment for the company's current requirement. The next stepping stone is appropriately mapping that into the core systems that require the data. This is applicable to client, broker and security reference data.

Marc Rubenfeld, head of Eagle Solutions, EMEA/APAC, Eagle Investment Systems: Given all of the considerations in sourcing data, it can be hard to identify just one as the single biggest issue facing financial services companies. Reconciling data from multiple sources can be a huge task, particularly as data volumes grow and you have to deal with duplicative and contradictory data or the challenges of reprocessing all of this information.

The bigger issue for sourcing data is whether you actually control it. To truly leverage your data, it's critical to have it centralized and integrated within the larger ecosystem so that institutions can go back and access the

data in a meaningful way. This allows you to enrich the data, repurpose it, or find new ways to apply it for reporting, benchmarking or analysis. This is critical as institutions look to optimize their investments in data management, to support or drive decision-making in the front office. It's important to realize that these data management projects are investments and that organizations are paying for this data, so they should have complete control of it and truly own it.

The costs of maintenance and the pressure so many institutions face with capital and resource allocation are causing a real trend toward managed services in data management. It's not just the available efficiencies; it's a data quality issue, as institutions rely on committed partners whose core job it is to manage data and, at the same time, enjoy the economies of scale to stay on top of it, as the inputs continually change.

Chris Johnson, head of product management, market data services, HSBC Securities Services: The biggest issue is the high cost of procuring data and retaining resources to manage it effectively. The greatest challenges are, first, defining the data content that meets the needs of the consumers; second, creating a governance structure and opera-



Marc Rubenfeld,
Head of Eagle Solutions,
EMEA/APAC
Eagle Investment Systems
T +44 207 163 3719
mrubenfeld@eagleinvsys.com

tional process capable of providing the right data on time; third, implementing IT systems that can support the operating model with control, lineage and flexibility; and fourth, procuring the data at a reasonable cost and with licensing permissions that enable the consumers to achieve their business purposes.

"To truly leverage your data, it's critical to have it centralized and integrated within the larger ecosystem so that institutions can go back and access the data in a meaningful way"

Marc Rubenfeld, Eagle Investment Systems

Virtual Roundtable

Dominique Tanner, head of business development, SIX Financial Information: Data sourcing is a classic case of cost versus benefit, where the benefit is improved data quality. The more resources you deploy, the better your data quality becomes, but this entails a cost increase. Sourcing the same data from multiple sources and scrubbing it can lead to the detection of more errors and hence better data quality overall. However, applying this strategy across the board implies sourcing costs that most firms can't or won't want to bear.

Adopting a differentiated data-sourcing strategy can help in achieving a better relationship between cost and quality. Not all data is equally important or has the same impact on business operations. This differentiation can be done by asset classes and markets, but also by data sets or individual data points. A multi-source strategy should be applied where it matters most and the risks are high, while a single-source or specialist source strategy may be sufficient for the rest. It is also vital to limit the overall number of sources in order to control complexity.

How does the way sourcing is achieved, or what the sources are, affect the quality that users will get in the end?



Financial Information



Dominique Tanner,
Head of Business
Development,
SIX Financial Information
+41 58 399 5577
six-financial-information.com

Harkins: Many organizations source the data around a specific business need or regulatory requirements. Also, organizational changes, business acquisitions, platform consolidation or regulatory reporting changes directly impact how sourcing is achieved. Finally, the industry sources that are utilized or the quality of the data housed in the system from conversions can directly impact the quality of data.

Rubenfeld: We define data quality by four measures: completeness, accuracy, consistency and timeliness. Traditionally, your data is only as good as your sources. If your data comes from an inferior or incomplete source, you will get low-quality data. If your data is sourced in an inefficient way, or if you have to reprocess the data, you're leaving yourself

> continued on page 12

EUROPEAN REGULATION ROADSHOW 2015

hosted by: **waters** **Inside Market Data** **Inside Reference Data**



DON'T MISS THE
UPCOMING EUROPEAN
REGULATION
ROADSHOW EVENTS!



#ERR

www.europeanregulationroadshow.com

HIGHLIGHTS INCLUDE:

- Reviewing the current European regulatory landscape: AIFMD, FATCA, Solvency II, BCBS 239 and MIFID II
- Understanding the regulatory drivers including transparency, capital requirements, systemic risk oversight and surveillance
- Bridging the gap between the regulatory drivers and business drivers: Balancing cost vs. regulatory compliance
- Is there sufficient cross-over between regulations to warrant a central regulatory data warehouse and reporting solution?
- In what ways can firms collaborate to reduce the overheads associated with the common data needed for regulatory reporting, and do the regulators need to help with this?

FINAL ROADSHOWS COMING UP:



ZURICH
Regulation Roadshow
27th October 2015



AMSTERDAM
Regulation Roadshow
26th November 2015



An innovative media event
powered by

Lead sponsor:
**BLOOMBERG
VIA**

Sponsors:

FACTSET



Interactive Data



**S&P
CAPITAL IQ**
BY MORGAN MULLER FINANCIAL



FINANCIAL INFORMATION

tullett prebon
Sponsor

Virtual Roundtable

> continued from page 10

open to errors or issues every time that transfer of information occurs.

With the right tools in place, you can optimize your data sources to get the most from them. For instance, you might have two data sources that by themselves are incomplete or inconsistent. With the right technology, you can compare and consolidate the data and create a hierarchy that maps out primary and secondary sources. This not only fills in the gaps, but also serves to validate the data to ensure end-users are receiving the highest-quality components every time. This is also why it's so critical to work with vendors that are all part of the same ecosystem, whose products and services can easily be integrated onto the same platform. It allows you to streamline your data management processes and cut costs.

"Adopting a differentiated data-sourcing strategy can help to achieve a better relationship between cost and quality. Not all data is equally important or has the same impact on business operations"

Dominique Tanner, SIX Financial Information

I would just add that the quality of data is often subject to how it's treated, processed and made available across the organization after it's been received from the original source. This is one of the key reasons it's so important to have a data governance policy in place, as it creates standards and expectations that lend to consistency and timeliness of data.

Johnson: If the sourcing is effective, then the data quality largely looks after itself and the need for repair is minimized. However, this presumes that all of the required data is available. In reality, there are many data gaps and these should be managed proactively with resolution pushed upstream rather than repaired in-house. This area is difficult and time-consuming and is often overlooked.

Tanner: Choosing the right data source or the right combination of data sources is a major factor in achieving high data quality. Not every data source will be able to provide the same level of quality through all data subsets it can offer. It is important to understand what the strengths and weaknesses of each source are and compare them with the data quality requirements.

A first, vital part is to know where the data source itself is getting the data from. At SIX Financial Information,

we aim to use primary sources whenever possible. Primary sources are the originators of the data, such as stock exchanges, depositories, lead managers or registrars. This leaves less chance for errors to accumulate in the distribution chain than if data is taken from another data aggregator (a secondary source).

The second step is to understand the data quality assurance checks and processes of the data source, as there is no point in replicating pre-existing checks at the receiving end.

Does automation of data processing affect data quality? Does automation make it more difficult to catch and correct errors?

Harkins: The quality of the data after automation depends on the amount of time spent analyzing and profiling the data and uncovering the true data content prior to sourcing it. Automation directly impacts and affects the data quality unless there are data rules and thresholds. In many high-volume systems, proofing and reconciliation is a daily routine and aids in the detection of errors.

Rubenfeld: We've found that automation is critically important to data quality, as it helps to discover errors that wouldn't otherwise be discernable and ensures

that they're flagged and corrected for the future.

Automation effectively creates a filter that ensures pre-established standards are met. This is important because data quality relies on absolute alignment, particularly for large, sprawling enterprises spread out across geographies and languages. Even the smallest discrepancies in interpretations around business ontology can lead to inaccurate and inconsistent data. This becomes more acute for complex securities such as fixed income or derivatives. Automation eliminates the "data silos" that are often at the root of data inconsistencies across organizations.

Beyond creating a data quality standard, another benefit of automation is that it allows you to establish an optimization program that cuts across all areas of the technology stack. This means you can apply analytics and instrumentation to manage third-party tools and measure their effectiveness. You can drill down into the operating data and, if you see that 20% of the failures are attributable to one specific vendor, you can take the necessary actions. Alternatively, it could be a particular portfolio that is proving to be more costly than the others to main-



*Amy Harkins,
BNY Mellon*

Virtual Roundtable

tain. With automation, it's easier to pull the necessary information and quantify it. Then, armed with the facts, you can go back to the client and have a productive conversation about re-pricing.

Johnson: Automation enables fast validation checks to be performed so that any quality exceptions that are identified can be resolved efficiently. The trick is to implement the right validation rules so that errors can be detected. This is the most sophisticated and critical area to get right.

Tanner: The way data is processed from sources, validated and disseminated to downstream applications has a significant impact on data quality. The good news is that data processing deficiencies are systematic in nature and, if they are corrected, they are not likely to appear again. The bad news, however, is that they are not easy to localize in today's complex IT environments.

Most of the deficiencies, and therefore perceived data quality issues, are caused either by misinterpretation of source data or data changes that were missed. It is vital to properly understand how a data source, whether external or internal, delivers the data and exactly what facts it represents before mapping them to the corresponding target data fields. In some

cases this might be straightforward, but in others it involves applying more complex mapping rules. If this mapping process is done over several stages, from inbound datafeed to a data repository and then onward to a consuming application, it makes it harder to find out what has gone wrong and where the error has occurred.

Are there better ways to process data that can raise its quality?

Harkins: An organization should have a centralized team that understands the business line needs, purpose and content of the data. This team should have standard level agreements or descriptions with the business lines they support, which clearly define the servicing of the data, the quality results and the reporting around the data.

Rubenfeld: Absolutely. It's vital to begin with a data warehouse that centralizes security master data, prices, issuers,

"An organization should have a centralized team that understands the business line needs, purpose and content of the data"

Amy Harkins, BNY Mellon

positions and transactions to composites, entity relationships, benchmarks and everything else. Create a data hub that incorporates all of the different sources of data and centralizes the information, down to asset-mix policies and business ontology standards. This is the only way to ensure that the entire enterprise is executing from the same playbook and has access to the same data.

The biggest factor for data quality is the flexibility in a data-centric model that delivers the same validated and enriched data across the back-, middle- and front-office functions. Data-centricity also allows for true transparency into the data lineage, from source to consumption.

When issues occur, users can go through the data and permanently resolve any inconsistencies. Flexibility is also critical as organizations seek growth. A data-centric model makes it exponentially easier to add asset types or target new regions, while a piecemeal approach of disparate systems and processes opens the door for all of the issues we've discussed around data quality.

Also, greater access allows the same data to be used for cash management, reconciliations and an accounting book of record. That supports the addition of new functions that help form an IBOR or PBOR [investment or performance book of record], including data enrichment,

look-through and exposure analysis, post-trade compliance and enterprise reporting, performance measurement and attribution, and *ex post* and *ex ante* risk analysis.



Chris Johnson,
HSBC

Johnson: The utopian solution is for standardized and consistent “single version of the truth” data to be supplied by all data vendors and for data gaps to be solved upstream at source. This is potentially achievable for commonly used data fields if all investment firms and regulators are prepared to sponsor and support this long-term and momentous aim. It is possible that the recently announced European Securities and Markets Authority instrument database could provide a template. Such standardization would also provide an opening for utilities to supply commonly used data with the possibility of long term savings for the industry.

Tanner: As mentioned previously, the key to improving data quality is to understand how the source data needs to be interpreted and what conventions and standards are used, as misinterpretation can lead to data quality issues on the receiving end. SIX Financial Information is aware of this and provides extensive

Virtual Roundtable

support in understanding and mapping data correctly, through in-depth documentation and the expertise of its customer-focused data consultants.

Getting this right in the first place is important, but the data interfaces and processing rules also need to be maintained over time. In today's world, changes happen frequently and ever faster. A key factor is choosing data sourcing partners with a reliable and predictable change management process that bundle changes into data feed releases and give customers enough time to make the necessary changes at their end and test any new features accordingly.

Do high expectations for data quality make it more or less difficult to manage sourcing, automation and processing effectively?

Harkins: This truly depends on the type of data being considered. Sourcing for security master file or broker data is more difficult to manage because vendors do not want to be responsible for the data. In many high-volume systems, proofing and reconciliation is routinely performed daily on aggregated transactions, potentially making account-level issues more difficult to identify.

Rubenfeld: You need high expectations to run an effective data management

program. Data quality projects can seem incredibly difficult if you don't have the right tools and partners in place. But you have to realize that you're building an ecosystem that feeds off itself, so if veracity and rigor are built into the processes and you have a fully integrated system, most of the work to ensure data quality occurs up front. If this attention to quality isn't built into the system, or if you're trying to piece together a program and don't have the necessary components to cover the full breadth of what you need, then it's going to be difficult to meet any expectations around quality.

Those who take an enterprise data management approach often just think about the operational data store, which integrates and validates data from different sources. They overlook other components that are critical to data quality and supporting a comprehensive

"Automation enables fast validation checks to be performed so that any quality exceptions that are identified can be resolved efficiently. The trick is to implement the right validation rules so that errors can be detected"

Chris Johnson, HSBC

data management program with business benefits. If users expect to employ their data to support risk management or performance measurement and attribution, they need more granularity and detail than is provided through a traditional operational data store. IBOR users need transparency for look-through and exposure analysis. Meeting these expectations requires an integrated and comprehensive program, extending from warehousing and validation of data to enrichment and access considerations. Data management initiatives should establish a firm foundation as a base to build an ecosystem that functions to achieve clear milestones and goals.

True data quality comes from managing the whole ecosystem. Once a comprehensive and all-encompassing program is in place, you should not only have high expectations around data quality, but also be confident that those expectations can be met.

Johnson: High expectations are extremely helpful in achieving data quality. If the consumer cares about the data and is prepared to shout about it, then the likelihood of success is increased. High expectations do, however, serve to raise the bar for quality validation, which increases the complexity and cost of data sources and expertise levels required to manage them. This is particularly relevant where

there is no official market data standard. An example of this is bond prices, where different prices are often available for a given bond that can be appropriate for different purposes. In summary, much complexity stems from the root cause that the concept of golden copy data only applies to a small proportion of data content at present. The existing data fabric across the industry therefore has variations in content and format that need to be supported.

Tanner: In principle, it is good to have high expectations for data quality. It sets an ambitious target to aim for that drives the underlying efforts to achieve it. Getting there will take time and resources, and compromises will need to be made in order to adapt to business reality. Identifying the areas with the biggest leverage is a key step in the process. Firms need to understand the required levels of quality in each data subset. The question is, what impact does a data error have and what are the associated risks, such as a settlement failure or a fine from a regulator? This needs to be mapped out holistically and corresponding priorities need to be set.

Achieving the required data quality levels involves deepening the partnership with data sources, technology service partners as well as helping to drive industry-wide initiatives to harmonize data and workflows.

Clearing the Road to Quality

Data quality cannot be addressed with a broad governance sweep and should be seen from consumers' point of view, says **Roberto Maranca**, managing director for enterprise data at GE Capital



Roberto Maranca

Are data governance plans proving effective for improving data quality?

In its essence, data governance paves the way towards a change of behavior and a care for data that should yield better data quality. The challenges are centered around its prioritization and pervasiveness. Many firms are planning too aggressively—trying to gulp the entire breadth of data in one go with policies, edicts and compliance programs. Data governance is a huge change in the collective company behavior and should be tailored to the company so that it is progressive, but most importantly sustainable.

What is the most important element for addressing data quality, and why?

Data quality must be seen through the eyes of the data consumers. Too often, “good data quality” is discussed as an absolute value. Therefore quality-level

agreements between data owners and data consumers should be brought about to detail what good looks like, in order to manage expectations and use company resources in the right direction. That should also simplify data quality communication, as an unqualified ocean of red, amber and green indicators can affect the confidence of those consuming the data.

Is automation having a negative effect on data quality?

Potentially, if automation increases the output of a process in terms of unit produced per time, and as a benefit is supposed to reduce the human touch-time in the whole process (e.g. risk auto-approval), there is a risk that a quality issue could be replicated on larger scale faster, and have a deeper impact. The solution is to design automation with more robust and tested quality control processes from the start.

Leverage Versatility. Deliver Stability.

Innovative technology solutions to help investment managers respond quickly and efficiently.

Your business demands a stable and flexible platform that enables you to consolidate accurate information from multiple global locations and business lines.

From a single platform, Eagle can support data management, investment accounting, and performance measurement and attribution requirements to help you more efficiently manage your investment portfolios.

[Enterprise Data Management](#) | [Investment Accounting](#) | [Performance Measurement and Attribution](#)

Contact us today to learn how Eagle can help your business:

Americas	+1 781 943 2200
Asia Pacific	+65 6432 0581
Europe	+44 20 7163 3827
Mainland China	+86 10 58771605



BNY MELLON

www.eagleinvsys.com

Eagle Investment Systems LLC, a leading provider of financial services technology, is a division within BNY Mellon Asset Servicing and is a wholly owned subsidiary of The Bank of New York Mellon Corporation.

©2015 Eagle Investment Systems LLC. All rights reserved.

EAGLE
INVESTMENT SYSTEMS



Incisive Media 28–29 Haymarket, London SW1Y 4RX