A new approach to data quality

Despite many years of investment in technology solutions and governance processes, the data quality problem has not yet been solved. Indeed, it has become such an issue that regulators are now requiring governance processes and internal controls to address data quality and consistency, particularly where it relates to risk management.

BCBS 239 Principle 3, 36 (a) states:

*Controls surrounding risk data should be as robust as those applicable to accounting data.*

The traditional approach

The traditional approach to improving data quality involves implementing a central repository that serves as the master source for reference data. This is either a commercial offering from vendors such as Golden Source, Asset Control, Markit, Bloomberg, etc. or internal development of a data hub or data warehouse. While these approaches are undoubtedly of benefit to almost any firm, they have some significant shortcomings – it is rare to complete 100% integration with the master source, there is often a lack of feedback of updates from business systems back to the source, it is hard to measure the quality of reference data in each system and track it over time, the implementation lead-time is often too long to cater for new regulatory demands and business requirements and, finally, the individual system owners are often reluctant to impose new business processes on their users.

The industry has long debated the merits of in-house development versus vendor solutions, and the relative flexibility of data hubs versus the acquired structures of data warehouses. There is, however, an alternative approach to the traditional options...

A new approach

Technologies such as big data, social networking, distributed databases, metadata and distributed ledger technology have provided infrastructure that can be leveraged to address the data quality problem. A new approach using these techniques can be implemented in a fraction of the time taken to implement traditional solutions and has a return on investment profile that meets the exacting demands faced in the current climate.

This approach is a new take on the utility model, one that doesn’t require wholesale changes to vendor feeds, or require implementation of new EDM platforms or to radically overhaul business processes. An approach that yields quick, measurable results but also fosters buy-in from individual system owners and other stakeholders.

Sample Screenshots

*Agree100™ quality scores for 5 sample systems*  
*Example field-by-field data exceptions*
The Investment Data Utility

The Investment Data Utility is a new organisation conceived specifically to address this challenge. It is part of a family of data quality solutions, each customised to suit the exact nature of the target customer, with parallel solutions available for Investment Managers, Hedge Funds and Insurance customers.

Based around a proprietary scoring model called Agree100™, The Investment Data Utility platform enables immediate comparison of reference data at the point of use from an unlimited number of internal systems. A series of consensus algorithms automatically mediate between the different sets of reported data, with structured workflows allowing challenges to the consensus. Any sets of data can be compared using the platform, with a metadata structure allowing each customers to choose which data is compared. Customisable algorithms support advanced consensus models and variable tolerances. The results provide quantitative measurement of the quality of data in use at the ‘coalface’, both on a daily basis and over time. More importantly, the analysis provides field-by-field lists of suspect data items, consumable in XML format for automatic repair or via filterable on-screen views for manual assessment. The platform can also source data from the point of entry for distribution directly to other systems or integrated with existing EDM technologies.

Enterprise and Community Deployment

The platform can be deployed as the Enterprise Edition for a single firm, mediating between the plethora of internal systems, or in Community Edition for a group of like-minded organisations that see the benefits of collaborative working. It is common to start with an Enterprise deployment to generate immediate benefits, then consider the Community model at a later date.

Business Benefits

- Rapidly quantify the data quality problem
- Support measurable change during data governance programmes
- Reduce the cost of manually repairing incorrect data
- Minimise operational risks arising from incorrect data
- Respond quickly to regulatory requirements

Key Functionality

- Quantify – immediate measurement of data quality across many systems
- Detection – identify suspect data items ahead of costly mistakes
- Repair – automated or manual repair of suspect data
- Collaboration – fix data once and distribute to other systems
- Sourcing – collect new data from point of creation and distribute across the firm
- Mediation – address differences of opinion – deliberate business need or data quality error?

To discuss your requirements or a pilot project, please contact us using the details below...