ASSET
CONTROL

BRICKENDON
transformational consulting

Big Data and Data Quality

Setting the Agenda for CDOs

Co-authored by
Asset Control and Brickendon

Industry Perspective

Data is everywhere and takes on a variety of forms. Whether it's facts, opinions, figures, statistics, charts or words, there is no doubting its importance in today's highly competitive business environment. Managed correctly, data has the potential to transform your business and be a source of sustainable competitive advantage. Managed incorrectly, it can become a liability, not only costing you lots of money but also your reputation. Today, every organisation is in the information management business.

In addition to data being everywhere, it is also constantly changing. As a result of the constant increase in the variety of data sources (e.g. alternative data and granular market data), the velocity of processing (e.g. real-time reporting and faster calculations) and the volume of data produced (e.g. social media, post-trade reporting), the data infrastructure has become more complex. Business processes are becoming increasingly data intensive and business users increasingly data hungry. Lineage and audit processes are also being increasingly scrutinized. As a result of these changes, a lot of strain is put on the existing data aggregation/collection processes, making it more and more difficult to ensure the data is relevant and customer-focused.

The above issues are magnified in a big data environment as the data becomes less structured and expectations about what can be achieved increase. Failing to manage your data correctly can quickly change it from an asset to a liability due to the costs of acquiring, storing and maintaining it. Irrespective of the enhancements to data analytics tools and systems, the basic principle of garbage-in and garbage-out still stands true. In fact, due to the proliferation of data analytics tools, poor data quality can lead to firms making incorrect decisions that can potentially harm their business.

In the same way that data and its uses have evolved, and still are evolving, so is the role of the Chief Data Office (CDO): a team/individual responsible for enterprise-wide governance and utilisation of information, in other words, data as an asset.
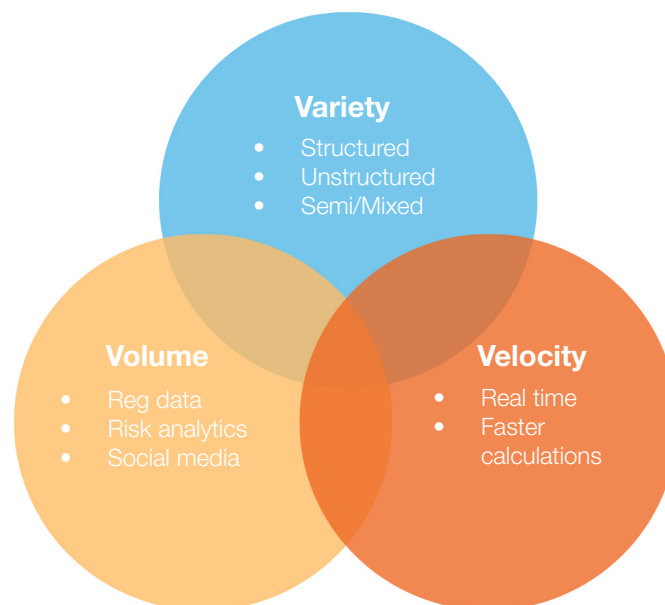

Variety
- Structured
- Unstructured
- Semi/Mixed

Volume
- Reg data
- Risk analytics
- Social media

Velocity
- Real time
- Faster calculations

Figure 1 - The 3Vs of Big Data

ASSET CONTROL

BRICKENDON
transformational consulting

Data Quality Drivers

Historically, the term "data quality" has been used to describe quantitative measures of timeliness, accuracy and completeness of data, all of which are referenced in various financial regulations such as Solvency II and G20 reporting. However, the main drivers for data quality continue to evolve:

- **Regulation:** Regulators are placing increased emphasis on lineage, both horizontal and vertical, following the data from source to target system, and auditing the impact on calculations and transformations. Data owners must now be able to reverse engineer, or replay, particular events (e.g. trade reconstruction in MIFID II) in order to remain compliant.
- **Global reach:** As businesses become increasingly cross-functional and global, there is a need to connect different data sets in order to gain additional insight or a fuller market picture.
- **Trust:** Trusting your data is important. Historically, the lack of transparency has undermined this confidence, preventing organisations from evolving their operations and integrating analytics to lower their costs and reduce operational risk.
- **Cost effectiveness:** Centralising data sources lowers data acquisition efforts, avoids remediation effort due to data duplication and saves on multiple vendor licences for sourcing data.

Quality Measurement Strategy

Depending on the size and the complexity of each organisation, a common question that often arises is around the data measurement and display strategy. Typically there are two methods of data flows that can be accommodated: federated and big data.
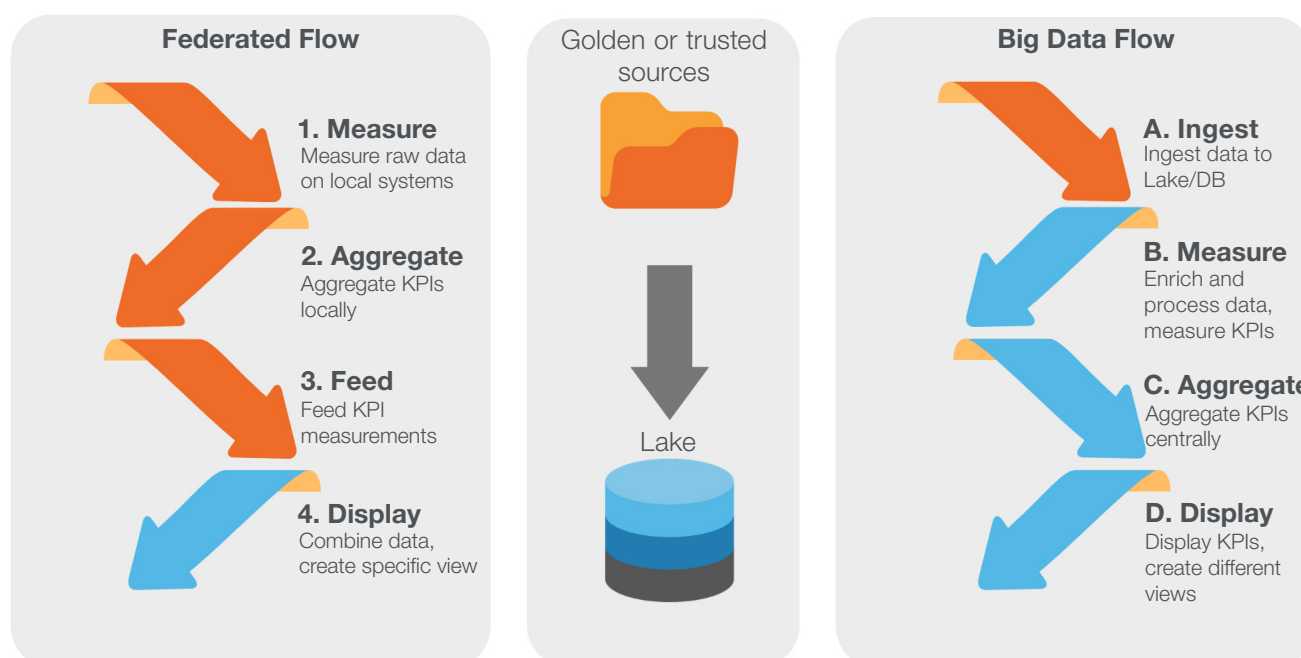

Federated Flow

1. Measure
Measure raw data on local systems

2. Aggregate
Aggregate KPIs locally

3. Feed
Feed KPI measurements

4. Display
Combine data, create specific view

Golden or trusted sources

Lake

Big Data Flow

A. Ingest
Ingest data to Lake/DB

B. Measure
Enrich and process data, measure KPIs

C. Aggregate
Aggregate KPIs centrally

D. Display
Display KPIs, create different views

Figure 2 – Federated vs Big Data Flows

ASSET CONTROL

BRICKENDON
transformational consulting

- **Federated Flow:**

Data quality measurements are taken at the source, aggregated locally to produce the required KPIs and then fed to the common data repository (Lake) to be displayed along with the KPIs of other trusted sources. The big advantage with this is the inherent domain expertise of the respective team. This not only improves the quality of the analysis, but also enables bespoke measurements and agile system changes to the respective data sets, ultimately speeding up the time to production. In a large enterprise environment with hundreds of systems, decoupling the delivery for each system will produce results more quickly and enable them to be acted on sooner. On the downside, a lack of consistency over the metrics definitions across different business units could impact the ability to combine data quality measurements in an accurate way. The identification and remediation of gaps will require further iterative manual work.

- **Big Data Flow:**

Raw data is fed into a central repository (such as a database, data lake or warehouse). The measurements are taken centrally for each attribute or CDE (Critical Data Element) across all systems, and then aggregated and displayed accordingly by a visualisation layer on top of the data repository. The benefit here is that there is a central repository designed for consistency and scaling up of the data quality measurements, with the only limit being the pace at which each system's data is ingested into the repository. If managed properly, this approach, despite carrying the overhead of a large scale change programme, will enable the creation and fine tuning of an overarching view of the business KPIs across all systems, with fairly low maintenance overhead once a system is onboarded.

A high-level review of the two approaches with the pros and cons is shown below:

Federated Flow Summary	
+	−
• Quicker time to production • Readily available domain expertise • Ability to use bespoke metrics	• Lack of consistency in definitions and results • Increased cost due to duplication of effort as number of sources increase • Limited ability to give the full picture

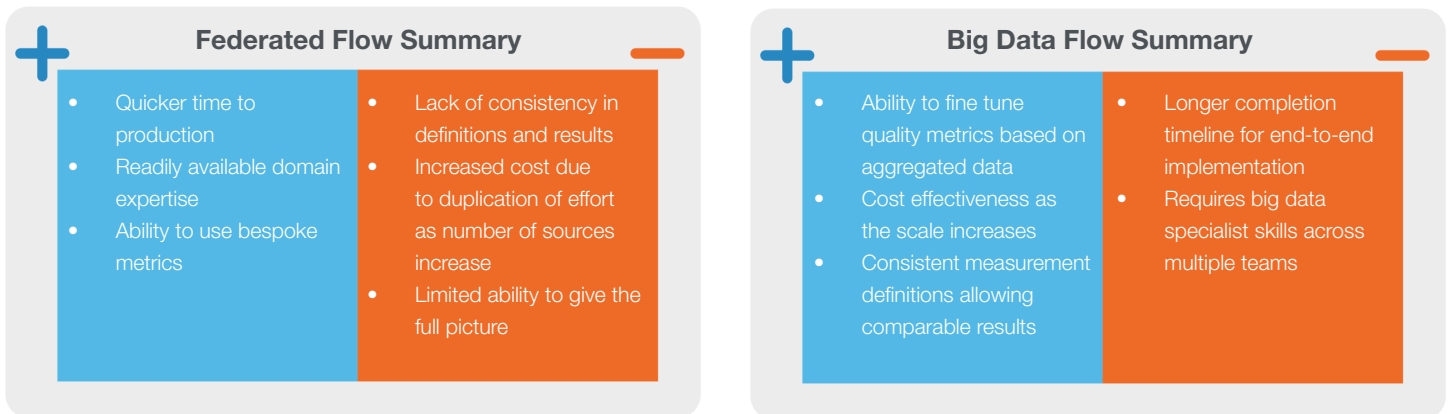| Big Data Flow Summary	
+	−
• Ability to fine tune quality metrics based on aggregated data • Cost effectiveness as the scale increases • Consistent measurement definitions allowing comparable results	• Longer completion timeline for end-to-end implementation • Requires big data specialist skills across multiple teams

Figure 3 - Comparative View

CDOs and business leaders should consider a variety of factors (e.g. time-to-market and adaptability of data sources) when making the decision on the measurement methodology. There will be cases where it is more beneficial to go for a hybrid model, that is to tactically implement the federated flow on a prioritised list of local systems, and at the same time plan for and/or initiate the strategic implementation for the big data approach.

The Role of the CDO Function

Many firms have appointed Chief Data Officers to address these challenges, improve data management practices and maintain a solid grasp on data quality across their data estate. This trend started with large, globally active firms with particularly convoluted data landscapes. However, the rise in data access and reporting requirements, combined with the opportunity presented by adopting new techniques has caused the number of CDOs to spread from global tier 1 firms to companies of all sizes.

Typically, firms will appoint a CDO for **two primary reasons:**
1. As a leader who will take charge and guide the organisation out of the data swamp.
2. As the focal point for operational change throughout the organisation.

The CDO task cannot be fulfilled by one person alone. As a result, we often see the emergence of a CDO office, which includes:

Product Owner

Responsible for data domains, including real-time market data, client data and external reporting needs

Individuals who delve deeper into the functionality with the aim of changing data culture as much as tooling

Data Stewards

CDO Change Managers

Focus on streaming the capture and processing of internal and external requirements to produce an effective data change programme

Figure 4 – The Role of the CDO

In other words, the remit of the CDO, either as an individual or an office (for the larger enterprise), has expanded significantly to become an organizational function of its own. It is now a living and breathing data-centric organisation that acts both as an advisory unit to the other divisions of the business, as well as funding and running its own data-quality infrastructure programmes.

What is Needed?

In order to be successful, a CDO (or CDO Office) needs to manage a number of tasks that are detailed below:

EXPLOITATION
Enable the business to make the most of the data within the organisation by implementing policies and practices that add structure, tidy up the data landscape and reduce exposure to regulatory risk due to data failures.

Items that should be included in the toolbox of the ambitious yet pragmatic CDO are:
- an agenda to build capabilities such as data lineage and bitemporality
- a process to ingest additional sources and onboard new use cases
- transparency on metadata through clearly exposing sources, quality metrics and usage
- easy access and enterprise search capabilities
- a strategy to scale data management.

POSSIBILITIES
Help the organisation navigate the data providers and fintech landscape. One thing that is often overlooked, and where a CDO should play a major role, is the world of opportunity provided by the fintech sector. There is currently a very rich landscape of technology and service providers that are all vying for access to the financial services sector audience.

DEVELOPMENT
Provide an interoperable data and technology stack to facilitate the adoption of new technologies, either in-house or via partnerships. The data architecture needs to be sufficiently open to integrate with third-party products and services. Additional sourcing options for the organisation to increase its capacity for change can include managed data services for specific use cases or data categories.

To realize this, businesses need:
- easy handover points,
- modern and easy-to-use APIs to help with straightforward integration of new data sets into business users' workflows and
- clear contextual information that helps users determine to what extent to trust the data. This is where quality metrics come in. Completely overhauling the entire infrastructure is typically not feasible. Fortunately, cloud strategies, NoSQL technologies and managed data services provide various paths to augment infrastructure rather than replace it.

MANAGEMENT
Manage regulatory data risks and ease pressure on the firm. The common element in all financial services functions is the increasing dependency on data. A common element of regulatory change is the increased granularity and scrutiny of the processes behind the data provided to them. The first question business leaders and risk managers will ask a new CDO, is how he or she can help address this issue.

How can Brickendon Help?

Using our **_Advise_**, **_Change_**, **_Do_** approach, we will tailor a programme to transform your business using the latest data analytics advisory, automation and programme delivery techniques.
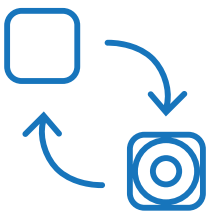
Advise

Create a data governance programme, which includes developing and executing strategies to mitigate data risk.

Understand the use cases of the data that is managed and develop an ongoing assessment of the data management capabilities.

Define KPIs and quality statistics that speak to the supported business processes.

Change

Evaluate data quality: Develop a process to continually evaluate the quality of the data and ensure that any defects are prioritised and addressed quickly.

Improve data quality: Establish or modify a programme specifically to improve data quality, including defining risks associated with the quality of the data.

Set unified control objectives: Rationalise and address issues such as data security and privacy through unified control objectives and activities that meet compliance or best practice requirements.

Facilitate use of analytics: Develop an analytics ecosystem that can capture, store and structure the data in a way that enables self-service analytics (where possible) and user interfaces to be used alongside big data and cloud infrastructure.

Do

Onboarding: Manage the onboarding of new users and data distribution to properly access data with approved tools.

Reconciliations: Prevent redundant reconciliations.

Standardisation: Simplify the number of data catalogues and glossaries to help with standardisation.

Data streams: Limit the number of independently arriving data streams that come from the same source.

Existing resources: Start small and look for sourcing options that can supplement existing infrastructure and solve specific business issues or investigate specific data sets.

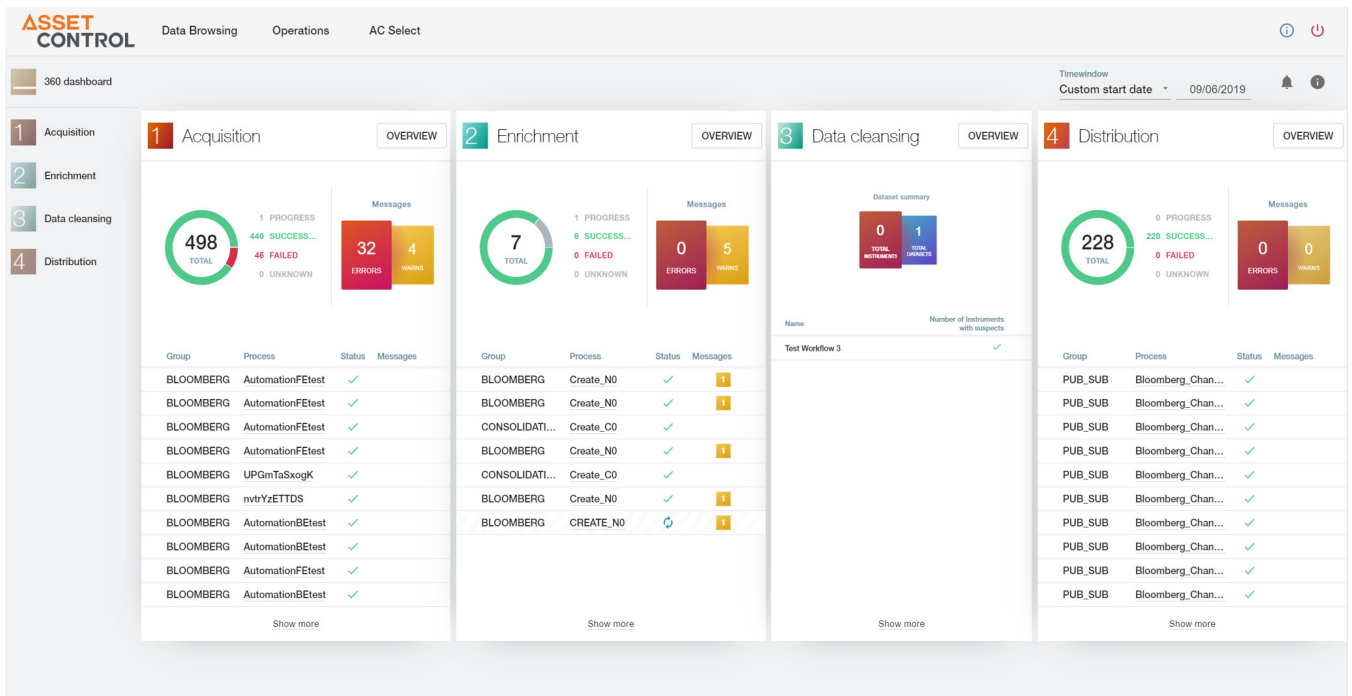ASSET CONTROL

BRICKENDON
transformational consulting

How can Asset Control help?

As a leading provider of standard software for financial data management, Asset Control provides its customers with solutions and infrastructure to source, master, derive and explore data assets. Asset Control's data services build and maintain data feed handlers, data models, mapping and cross-reference rules, data derivation and golden copy logic.

Financial data specialists and engineers work closely with all major data vendors to provide up-to-date loaders and data models. Clients don't need to monitor vendor change notifications, analyse changes, or implement and test loaders and models; they simply need to upgrade to the most recent maintenance release that is regularly provided. The close cooperation with many data vendors and our specialised knowledge ensures changes are incorporated efficiently, especially in the case of inconsistencies and/or anomalies which require close communication with the vendor.

By using Asset Control's data integration services, clients can onboard new data feeds more quickly, collate data from different sources into a single data model, be flexible in their sourcing strategy, reduce risk during implementation and maintenance, and work with the latest loaders and models for superior data management. This includes:

- Flexibility in sourcing and building master data for post-trade operations
- Easy access to the master data and configuration options to get the mastered data where it needs to be, quickly
- Large library of out-of-the-box loaders and mapping rules enabling cross-referencing and the cross-comparison of content
- Advanced analysis with vendor-by-vendor, field-by-field comparison of data formatting and content fields

About the Authors

Brickendon is an award-winning global transformational consultancy specialising in innovative solutions, technologies and products that disrupt the market and solve our clients' challenges quickly and efficiently. We do this in weeks and not months or years, saving them time, money and eliminating risk.

We are experts in digital, data and automation and use the latest data analytics, science, technology, machine-learning and artificial intelligence (AI) techniques to help you excel in your chosen market. We understand data and are here to help you fully realise the value and benefits of what you have and transform your raw data into a valuable business asset.

Find out more at www.brickendon.com

Contact details

T: +44 (0)203 693 2605
E: info@brickendon.com

instagram.com/brickendon_consulting
twitter.com/BrickendonIntl
linkedin.com/company/brickendon-consulting
facebook.com/BrickendonConsulting

Asset Control is the market leader in data quality software solutions for financial data. Focused on business user enablement, we help clients simplify complexity and ensure users across buy and sell side make the most of their data assets by providing easy data integration, data cleansing, distribution and data discovery solutions.

We service a blue-chip client base globally and our award-winning solutions provide rigorous processes to secure high-quality data, easy integration into business user workflows and a trusted environment for advanced analytics. Delivered through managed services, cloud or on-premise deployment, our highly scalable products help the world's most successful financial institutions meet their risk management, valuation, security master and operational needs with mission critical reliability.

Find out more at www.asset-control.com

Contact details

EMEA	T: +44 (0)20 7743 0320
North America	T: +1 212 445 1076
APAC	T: +65 6408 0750
General Enquiries	E: info@asset-control.com

twitter.com/asset_control
linkedin.com/company/asset-control