White paper

Proactive data governance:

Getting ahead of the game

Table of contents:

1. Introduction

Today's firm operates in a data-filled world – Master data, Reference data, Big data, Open data, Linked data. The modern business realizes the importance of data and is rightly enthused about it, seizing the rich business opportunities it offers.

However, for the majority of companies, management of this data hasn't followed in step. Essentially they don't know what they know, and, if they do, they don't know where to find it. If they successfully locate the data, they can't then be sure whether or not it is good enough to use.

Data Governance is a key data management discipline, but is often wrongly viewed as an IT function, distinct from the business. In fact it's now more important than ever for companies to adopt a strong, business-centric Data Governance approach, before the volume, variety and velocity of data increase even further.

While it is great that anyone correctly sees the value in implementing Data Governance, an IT led Data Governance initiative is likely to be more challenging and likely to fail. For several decades now, business users have lived under the misapprehension that their IT department owns the data of their organization. In reality IT own and are responsible for the infrastructure that the data is held on. What that data is, how and why it is held is subject to business requirements. For example, if the business wants to launch a new product, they would advise their IT department what details they want to hold about it on their systems. They do not tell IT that they are going to launch product X and to "please add some fields of your choosing to our product system to hold details about it", do they? While this may sound obvious, it is sometimes difficult for business users to take that next logical step and acknowledge that they own and are responsible for the data.

An IT led governance project will result in a system which may make sense for the IT department. They will know what data is held and where to find it. However as they are neither producers nor consumers of that data, it is unlikely that they will ever truly understand the business detail and context of that data. If others in the business know who in IT to ask, then they may find the information they and the firm are looking for. If they don't they won't, or they may unintentionally be pointed towards the wrong data. If this sounds familiar it's because this is the usual approach. Data is increasingly one of a firm's greatest assets and yet most employees don't even know what they have or where it's located and how it is used.

While many IT departments have recognized early on that data needs to be governed and have commenced Data Governance initiatives, their business users need to be on board and willing to take responsibility and accountability for the data, in order for it to be of use to the firm.

---

2. Definitions

It is more important than ever that businesses take ownership of and proactively manage their data. However, before we look at the reasons for such urgency, let's make sure we are all talking about the same thing.

'Data Governance' is a term increasingly being used but still means different things to different people. For the purposes of clarity it is best to make it clear what we mean by it and some related terms:

What is Data Governance?

Data Governance is simply a framework by which you can proactively manage your data in order to meet your business needs. At a minimum it will include:

- A policy to mandate how your organization is going to manage its data.
- Definitions of roles and responsibilities concerning data.
- Processes i.e. what to do to manage the data and indeed how to manage it.

It is critical to understand the different types of data when working within a Data Governance framework.

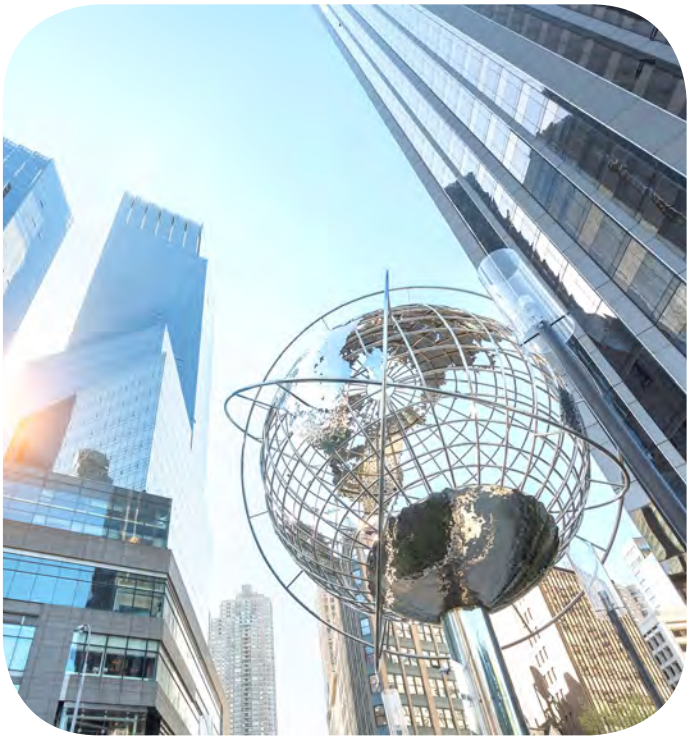
Master data is a single source of data that is critical for a business to operate and designed to be used across processes and systems. For example, customer, product, employees, suppliers etc.

Reference data is a type of data most often managed as master data, and refers to data used to categorize other data. It usually consists of lists of codes and corresponding descriptions or definitions e.g. ISO country codes.

Big data refers to data sets that are particularly large, often from new and unstructured and rapidly changing sources. Big data is commonly defined in terms of high volume, hypervelocity and high variety e.g. web logs, smart meter readings, social media feeds.

Open data is about sharing data in open non-proprietary formats. Examples of open formats include GIF, HTML and CSV. Open data is gathered or created by one organization and shared for reuse in a publicly specified format. It enables data mash ups (the creation of a website combining data from two complimentary sources, e.g. Google maps is often a source for data mash ups with links to other resources as appropriate, such as geo-coded photos from other websites) and supports the growth of the digital economy.

Linked data is a concept that transforms the web into a web of facts (instead of a web of documents), by connecting related data that was not previously linked. However if you don't understand the ever increasing sources of data that you link to, analyze and then make decisions on, can you be confident that you are making the right decisions? This is where Data Governance comes in.



What Data Governance is not

Data Protection: the use of policies and techniques to ensure the privacy of data i.e. that it is only seen or shared with those entitled to have access

Records Management: a proactive approach to managing records throughout their life cycle, from creation through to ultimate deletion

Data Retention: the policies and processes for ensuring that data is regularly archived and deleted in order to meet both legal and business requirements

These are all important data management disciplines in their own rights which need to be considered and implemented for your organization. However, they are NOT Data Governance.

What's the difference between Data Governance and Data Quality?

Data Governance is the framework you need to put in place to proactively manage your data. Linked to this is the presumption that your business needs the data being used to be of a certain standard in order to make the right decisions in running your business.

Data quality is measuring and monitoring how good certain sets of data are and improving the quality of it where it is necessary or appropriate to do so. This is accomplished using the processes, roles and responsibilities defined by your Data Governance Framework and will include things such as data quality reporting and data quality issue resolution.

Why do you need both Data Governance and Data Quality?

These two complimentary data management disciplines have a symbiotic relationship and one without the other is of little value.

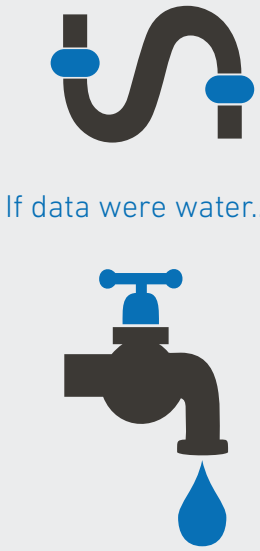
You could certainly set up data quality reporting without a Data Governance framework in place but to whom would you report? And who cares or is accountable/responsible for any remedial activity if the reports reveal that the data isn't up to standard?

Similarly, Data Governance on its own is of limited value. The purpose of having a Data Governance framework is to manage and improve data quality. Why then would you go to the effort of defining and implementing a framework, briefing numerous stakeholders and developing draft processes, if you were not going to use it to monitor and improve the quality of your data?

The ability to measure performance, review strategies and manage risk is damaged if you do not have good enough data available to support these activities.

Raw data is an extremely valuable asset. If not properly audited and used, however, that asset won't be realized.

3. Data Governance analogy



If data were water...

Data Quality
ensures water is pure and does not get contaminated

Data Governance
makes sure the right people with the right tools are responsible for the right parts of the water system

4. Why your business needs Data Governance

Failure to implement effective Data Governance impacts your competitiveness and compliance. This was highlighted in our 2017 global data management benchmark report:

84% of financial institutions see data as an integral part of forming a business strategy

74% of financial institutions believe that data quality issues impact customer trust and perception

50% of financial institutions say that they rely on educated guesses or gut feelings to make decisions based on their data.

The most common risks include:

- Non-compliance and regulatory penalties (56%)
- Downturns in customer loyalty and/or satisfaction (60%)
- Lost customer revenue (47%)

There are a number of business drivers which are actively supported by a solid approach to Data Governance as illustrated in the table below:

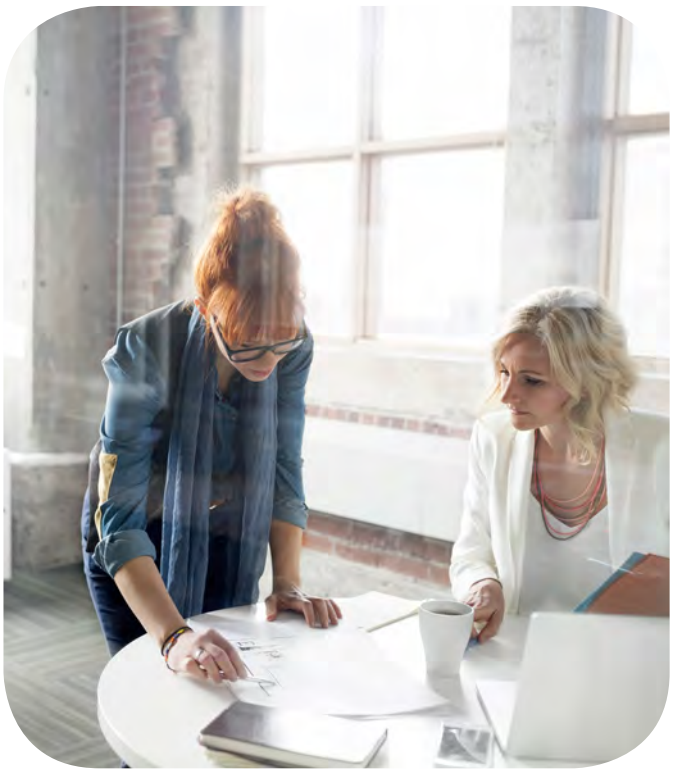
Business drivers for data governance

Business driver	Details	Example
Up-to-date data	Data is a valuable asset, but not if it's left to languish. If left unchecked the value of data as a business asset erodes and confidence in its accuracy and usability decreases.	Often after defining their requirements for the data to be held, business users do not define or implement processes to maintain and monitor the quality of that data. This results in 'data neglect' and a decreasing value of the data over time. In addition, many organizations do not have the ability to measure and prove the quality of their data. If business users have no confidence in the quality of data (rightly or wrongly) it leads to finding work-arounds to cleanse or supplement the data before they can use it. This not only adds time and cost to the processes, but also results in a proliferation of enduser computing solutions. Companies then have to deal with managing additional (and often unnecessary) databases held on Microsoft Access or Excel.
Compliance	Managing and mitigating compliance risks becomes harder and more expensive as the volume and variety of your data expands, growing increasingly more unwieldy.	Many organizations have, to date, chosen an approach of attempting to comply with the minimum requirements of whichever rule or regulation impacts them. Such a 'check list' approach to meeting regulatory requirements around Data Governance and data quality is flawed. It can lead to subsets of data being well controlled and managed (for example the data used in Solvency II calculations), but does very little to develop and embed the cultural change required for a company to adopt a proactive stance towards their data. Over the years it has also become apparent that the regulators' own understanding of both data quality and Data Governance is improving and their requirements are becoming more detailed and wide ranging with each new regulation.
Proactive compliance	Business benefits by doing more than the minimum required.	Businesses in regulated industries, such as Financial Services and Pharmaceuticals, have had to embrace Data Governance (albeit to a lesser or greater degree according to whether they take the aforementioned checklist approach or embrace a more beneficial holistic approach). The more proactive companies, in less regulated industries, have recognized that embracing Data Governance can give them the competitive edge. This is especially true of companies whose main commodity is or relies on data e.g. media companies.
Cost reduction	Operational inefficiencies persist and are often not even recognized.	How many teams in your organization devote considerable amounts of time to manual workarounds? A holistic approach means your organization spends considerably less time and money on ad hoc manual solutions for obtaining, reformatting, supplementing and cleansing the data required to support your processes.

Business drivers for data governance (cont...)

Business driver	Details	Example
Customer service	Poor data results in customers experiencing poor service, which compounds the longer it continues.	How often have you heard of or experienced bad customer service that you can attribute to poor data quality? This can include important mail being sent to the wrong address or inappropriate products being offered. Some are amusing like offering pre-approved car insurance to a bed-bound 94 old who has never held a driving license. However, some are more serious, for example county court judgments threatened (or even made) against individuals because payments from an insurance company have been sent to the wrong address. Not all poor customer service experiences stem from bad data, but it is surprising how high a proportion it can be, if you analyze your complaints database.
Changing data landscape	The exciting new opportunities with Big, Open and Linked Data present Data Governance and data quality challenges of their own.	The concept of owning data obviously has to change and key Data Governance concepts and processes need to be adapted to work for these sets of data. It is therefore critical that an organization has the solid foundation of a good Data Governance framework in place over the data they capture and create themselves. And implementing this before attempting to work out how to apply Data Governance for Big, Open and Linked Data or even using those data sets ungoverned. After all you can't hope to successfully combine large volumes of diverse data with your existing data, if you don't understand and manage the data that you already have.

Data is growing and diverging increasingly quickly – which results in an increasing urgency to implement a Data Governance framework. If you don't understand and control your data now, how can you hope to do with the current speed of change?



5. Where do you begin?

Dealing with the issue of how to govern your data requires a systematic, customized approach, both including and being led by the business. This means your firm's particular approach will be shaped not only by best practice Data Governance methodologies, but also by the reality of the business itself. Effective Data Governance means ensuring that data remains fit for purpose, relevant and useful.

Solid Data Governance should also look to optimize and automate the supporting processes wherever possible, making governing your data as easy and repeatable as possible. The removal of barriers to embed Data Governance practices is an important part of successful initiatives.

6. What do you need in place?

As with all things data related there are three main components.

People are the most critical component; it is vital that you define the roles and responsibilities that your organization will need in order to define and manage data. It is equally important to make sure that the roles and responsibilities are aligned with your organization. Do not insist on implementing best practice roles and responsibilities if they will not work for your organization. Look instead to adapt them to something that will function for you.

Secondly, you need to have some **processes** in place, to ensure that your newly appointed Data Owners and Data Stewards (or whatever you choose to call them) all manage their data in a consistent manner. Failure to have defined processes could mean that you end up with your data in a worse mess, with different parts of the organization applying different approaches and processes.

Finally, while Data Governance is primarily about the people and process, let's not forget the third part of data's holy trinity of people, processes and **technology**.

7. The value of data quality tools

Data quality tools can add significant value in supporting your Data Governance framework.

Tools on their own are not the answer, but you can successfully use tools to support and operationalize some of your processes and if you want to use bigger volumes of data then you need all the help you can get.

How data quality tools can help?

While people are familiar with the concept that their IT department uses tools to design and model databases and are also comfortable with the concept of data quality tools being those that measure and report on data quality, the tools available are wider than that and can support Data Governance in the following ways:

Data profiling tools:

These are a valuable resource in discovering vital information about the data you already hold. They can be used to determine the values held in fields and the relationships between them. Often businesses do not fully understand the breadth and content of the data they hold. You can't hope to govern that data if you don't know what data you have.

Glossary and metadata tools:

Having discovered information about your data, it is useful to have somewhere to store it for future reference and management. This can be supplemented with business definitions, reference data, lists of values and mappings between reference data codes held in different systems. An important part of Data Governance is defining what the data is and means. It also includes defining the business rules and data quality rules that define what makes data good enough for business use. Some tools include workflow capability. These can facilitate the processes for Data Owners and Data Stewards to review and approve data definitions and reference data values.

Data quality monitoring and reporting tools:

As mentioned these are probably the data quality tools which business users are the most familiar with and are used to report the standard of data held to the Data Owners and Data Stewards. This enables them to take action if required in order to investigate and resolve any highlighted issues.

8. The solution

Data is potentially an extremely valuable asset. Data governance has to be aligned with business goals in order to realize that worth. Success depends on building and implementing a holistic Data Governance framework that prompts and supports the use of business rules to regularly check data quality and measure performance against organizational KPIs.

Business comes first. Before leaping into setting up a Data Governance initiative you need to be sure that you are clear why you are doing it and what it helps your company achieve. Failure to align a Data Governance initiative to the strategy and business objectives of your organization will result in an effort which is not focused on the benefits which your company is seeking to attain. Any Data Governance project or program needs to help your company move towards its goals and deliver its strategy – if it doesn't do this, why are you doing it?

Good Data Governance is foremost about assigning the right people to the right roles and establishing good processes. These processes can then be well supported by utilizing the right technological tools. Such tools help operationalize processes and, as the data involved grows in diversity and sheer size, automation becomes ever more important.

Where automation is concerned comprehensive reference data is vital. Pairing technological tools with comprehensive reference data ensures automated processes draw on valid data at every stage.

9. Top tips to implement a Data Governance framework

1. Identify the priority business driver(s) your business has established and start your Data Governance work with a pilot on the data needed to support that business driver.
2. Bring senior business stakeholders on board as early as possible in your project.
3. Clearly define your Data Governance roles and responsibilities and assign the right people to the right roles early on in the project.
4. Use appropriate tools to support the processes you need to embed as business as usual.
5. Develop an extensive communication and training program to help business users understand and embrace the change required to embed a Data Governance framework.
6. Accept that you will have to evolve your Data Governance framework as the result of the pilot, the evolution of your business and the changing data landscape.

10. Summary

Data is a valuable asset for the modern organization. Data Governance allows you to realize that asset. The nature and volume of data available to businesses is changing at an incredible rate which means the potential costs of foregoing correct Data Governance are rising all the time. Don't put off until tomorrow, what you needed to do this morning.

Increasing regulation has made it more important than ever for companies to use Data Governance. The sensible approach is to embrace the changing regulatory environment instead of simply reacting to it. Even firms operating in less regulated sectors are becoming wise to the competitive edge that governing and understanding your data brings. Companies embracing a holistic approach to implementing Data Governance will enjoy more benefits as a result.

Data Governance initiatives have to involve and be led by the business and use a structured methodology to be successful, cost effective and most of all relevant!

When is the best time to get on top of and ahead of managing your data? How about now? Learn how we can help.

Get started

About the author

Nicola Askham, The Data Governance Coach, is an independent data management consultant. Her experience in coaching both regulatory and non-regulatory organisations to design and implement full Data Governance frameworks, is unique within the Data Governance field. The coaching approach enables organizations to self manage the process beyond initial implementation.

Nicola's coaching and Data Governance workshops, including Solvency II, ensures your Data Governance framework is embedded as an integral part of your business as usual policy. The benefit for you is that once the framework is in place, your organization will be confident, competent and compliant.

Nicola Askham
The Data Governance Coach
Independent Data Management Consultant

Experian
53 State Street
Boston, MA 02109
T: 1 888 727 8822
www.edq.com