ISO 8000 Part 150

A framework for data governance

A White Paper from DPA

Data and Process Advantage Limited
Advantage House, Stowe Court, Stowe Street, Lichfield, Staffordshire, WS13 6AQ
Tel: +44 1543 401144
URL: www.dpadvantage.co.uk

Table of Contents

1 ISO 8000:150–Overview

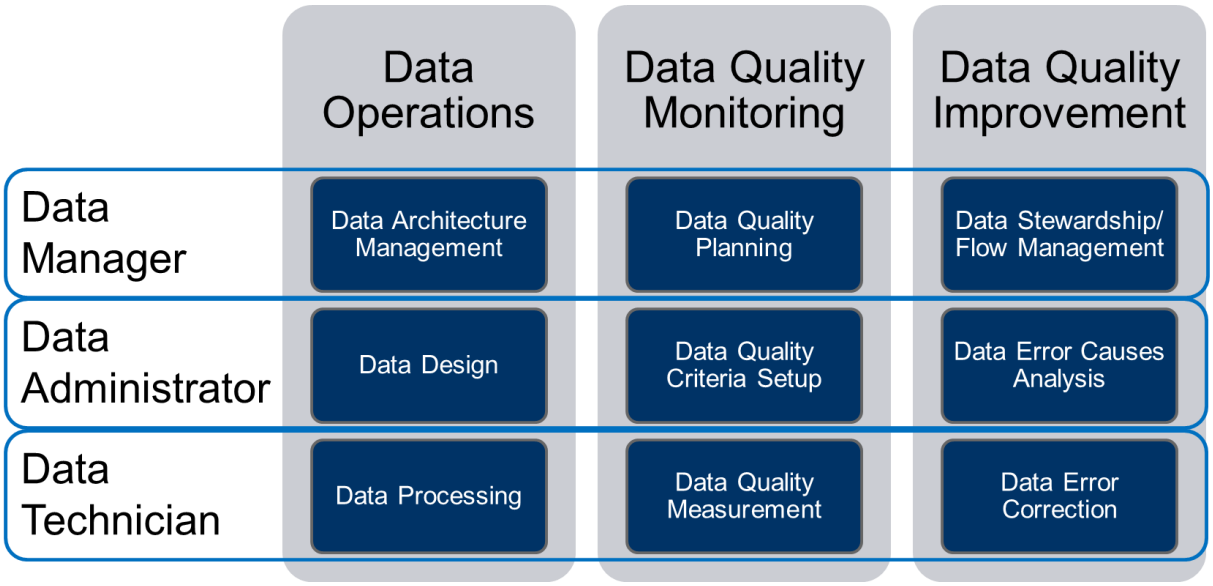
For many organisations, the challenge of managing data effectively as data volumes grow exponentially and demand for data continues to grow is a significant challenge. Fortunately, the ISO 8000 series of standards provides a range of definitive standards relating to data quality and ISO 8000:150 in particular provides a framework for data quality management. In this White Paper I will present an overview of the standard, look in more detail at specific areas and will consider the implications for data management and data governance generally.

The actual standard title is "Master data: Quality Management Framework", however, the approach is applicable to most data quality management contexts. There are a few fundamental principles that underpin the usage of the standard:

- **People** – data quality management is a people based activity and not a technology implementation
- **Process** – effective management is based upon a number of key processes
- **Continuous improvement –** as well as striving to continuously improve the quality of data, the processes used to achieve this should also be continually improved

These three principles are very important and, for me, the last principle is key in that it indicates that you do not need to establish large and complex processes and organisations before you can start to apply the approaches in this standard.

The methodology defined by the standard is summarised by the following nine box model:



This model is divided into three vertical 'processes' and three horizontal 'roles'.

The three key processes are as follows:

- **Data Operations** processes focus on the factors that affect data quality and the usage of data
  - **Data Architecture Management** manages the organisation wide data architecture
  - **Data Design** manages data standards and definitions, database and system implementation and configuration
  - **Data Processing** covers activities that create, modify, update and transfer data
- **Data Quality Monitoring** defines a systematic approach to assess the levels of data quality
  - **Data Quality Planning** sets the objectives of data quality management to align with organisational objectives
  - **Data Quality Criteria Setup** sets the measures and methods to assess data quality
  - **Data Quality Measurement** is the process that utilises these data quality criteria in order to assess data quality levels
- **Data Quality Improvement** process corrects data errors detected and eliminates the root causes of data errors
  - **Data Stewardship and Flow Management** is the process that analyses data flows and responsibilities and manages the stewardship of data operations
  - **Data Error Cause Analysis** is the process the identifies root causes of data errors in order to prevent them reoccurring
  - **Data Error Correction** is the process that corrects data that does not meet standards or data quality criteria

The standard also defines three generic roles – Data Manager, Data Administrator and Data Technician. For many organisations, these levels are probably too simplistic, however, they do provide an indication of whether the low level processes are strategic, tactical or operational.

2 ISO 8000:150–Data Operations

The Data Operations processes focus on the factors that affect data quality and the usage of data:

- Data Architecture Management manages the organisation wide data architecture

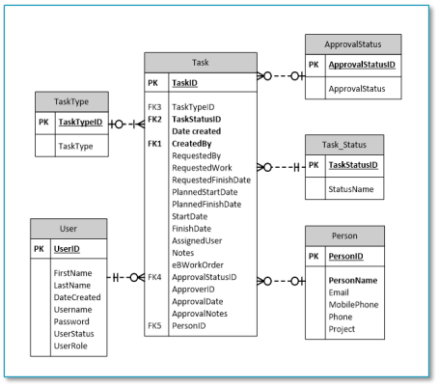
- Data Design manages data standards and definitions, database and system implementation and configuration
- Data Processing covers activities that create, modify, update and transfer data

Each of these can be considered as 'leading' indicators/ factors for data quality activities. We will look at each of these in turn…

2.1 Data Architecture Management

In most organisations, data is distributed across numerous data stores which means that data quality cannot be effectively managed if you do not approach data storage in a systematic way. This process includes activities such as:

- Management of organisation wide conceptual data models, perhaps as part of an overall Enterprise Architecture approach
- Management of organisation wide data standards including maintaining these as data requirements change
- Interfaces and coordination between data systems and stores
- Approaches to Master Data Management tools and approaches
- Accessibility and data security considerations



These activities also link to other data processes:

- Aligning Data Quality Planning to Data Architecture Management
- Links to Data Stewardship/Flow Management activities to understand data interactions across an organisation
- Input to Data Design activities to ensure that detailed data standards and conceptual models align to strategic ones

2.2 Data Design

Data quality errors can either arise from user errors or errors of data definition. The former can be corrected relatively easily, however, errors of data definition can be difficult to resolve, particularly if they have existed for some time. Key data design activities include:

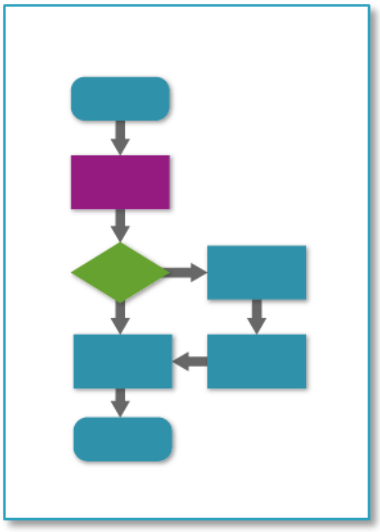
- The development of logical and physical data models
- Field level data standards and rules
- Awareness of data requirements arising from the overall Data Architecture approach
- Consulting other parties to understand the relationship of data to other systems/data stores and consulting users to ensure data standards meet data quality requirements
- Maintaining system configurations to align with these data standards

These activities also link to other data processes:

- Outputs of data design activities should influence Data Architecture Management
- Ensuring that Data Processing is undertaken in line with data designs
- Informing and being informed by the activity of creating data quality criteria

2.3 Data Processing

Data processing activities consider the actual processes of data update and creation. This is a key activity to manage effectively in order to get early warning that data quality problems are being created – using the 'dirty lake' analogy, this is like checking that the incoming water is clean. Activities in this process include:



- Monitoring data creation and update processes themselves
- Logging data usage and update times
- Awareness of the different roles undertaking data processing

These activities link to other data processes:

- Processing should be based upon and tested against Data Design outputs
- Results of data processing should be measured as part of Data Quality Measurement

Overall, these processes and activities should support organisations in setting a clearer framework for understanding and controlling the 'leading' indicators of data quality.

3 ISO 8000:150-Data Quality Monitoring

Data Quality Monitoring processes define a systematic approach to assess the levels of data quality:

- Data Quality Planning sets the objectives of data quality management to align with organisational objectives
- Data Quality Criteria Setup sets the measures and methods to assess data quality
- Data Quality Measurement is the process that utilises these data quality criteria in order to assess data quality levels

Unlike the Data Operations processes which focus on the 'leading' factors for data quality, these processes focus on the 'lagging' indicators i.e. identifying the level of quality.

3.1 Data Quality Planning

Whilst many people and organisations would ideally like 'perfect' data, the reality is that the time/ effort/ costs to achieve this are unrealistic. Therefore, the Data Quality Planning process sets the objectives for Data Quality Management to achieve overall organisational aims. This includes ensuring that there is a consistent approach to data quality supported by a detailed Data Quality Plan. Specific activities include:

- Agreeing and managing the organisational objectives for data quality based on internal and external requirements
- Managing the assurance processes for data quality management
- Planning of the activities to deliver the required level of data quality including specific tasks, timescales, resources and budgets
- Control of factors affecting data quality
- Gaining executive support for the Data Quality Plan

This activity also links to other data quality processes:

- Interface to the Data Architecture Management process to inform this process
- Results of the Data Quality Plan will also inform the Stewardship/Flow Management process
- Data quality planning provides the framework for the Data Quality Criteria Setup process

3.2 Data Quality Criteria Setup

In order to deliver the Data Quality Plan you need to be able to assess current levels of data quality. This involves the establishment of a number of Data Quality Criteria (which can also be known as Data Quality Rules) that details specific tests of the validity, completeness, uniqueness and accuracy of data. The two main activities of this process are:

- Identifying the data quality criteria, the target data and the measurement method
- Refining data quality criteria based upon the results of data quality measurement

This process links to some of the other processes in the framework:

- Criteria will be informed by the overall Data Quality Plan
- The agreed criteria are a key input to the Data Quality Measurement process
- Data Designs will also be a key input to the process of developing and agreeing data quality criteria

Each Data Quality Criteria should include:

- A description of the Data Quality Criteria
- What the Criteria represents
- The consequences of data failing the Criteria
- Status of the Criteria – e.g. Proposed, Regular, Ad Hoc
- Key stakeholders for the Criteria
- Relevant data source(s) and method of analysis
- Frequency of reporting
- Target quality levels by organisation and function

3.3 Data Quality Measurement

So far we have set the overall quality objectives for the organisation and have developed the supporting data quality criteria, however, we have not yet assessed the actual quality of the data!

The Data Quality Measurement process utilises the Data Quality Criteria to measure current quality which could be done manually or by specific tools.

There should also be suitable analysis and presentation of the data quality measurements in order to inform wider organisational processes and stakeholders.

This process is informed by the Data Quality Criteria Setup process and is a key input to both the Data Error Cause Analysis process and the Data Error Correction process.

4 ISO 8000:150-Data Quality Improvement

The Data Quality Improvement process corrects data errors detected and eliminates the root causes of data errors:

- Data Stewardship and Flow Management is the process that analyses data flows and responsibilities and manages the stewardship of data operations
- Data Error Cause Analysis is the process the identifies root causes of data errors in order to prevent them reoccurring
- Data Error Correction is the process that corrects data that does not meet standards or data quality criteria

Whilst the Data Operations processes focus on 'leading' factors for data quality and the Data Quality Monitoring processes are 'lagging' indicators of data quality problems, the Data Quality Improvement process should correct data to the levels intended.

4.1 Data Stewardship/ Flow Management

Due to the interactions of data with processes and the way that data can be shared across systems and data stores it is important to understand how data 'flows' across an organisation in order to be able to manage its quality effectively. The two core activities in this process are:

- Defining who is able to access and change data in various systems i.e. Stewardship assignment, and on-going management of these authorisations
- Data flow management is used to understand the relationships between the same data in different data stores and how changes in one should 'flow' to others. This is a component of Master Data Management and will typically involve coordination across a number of organisational units

These activities also link to a number of other data quality processes:

- There is a close link between Data Flow Management and Data Architecture Management since one defines the data relationships and the other how data flows between them. From a personal perspective, I would have seen these grouped as a single process in the standard
- This process is informed by data quality planning since this sets the framework and priorities for this activity
- Data Error Cause Analysis can take information from this process into the analysis process

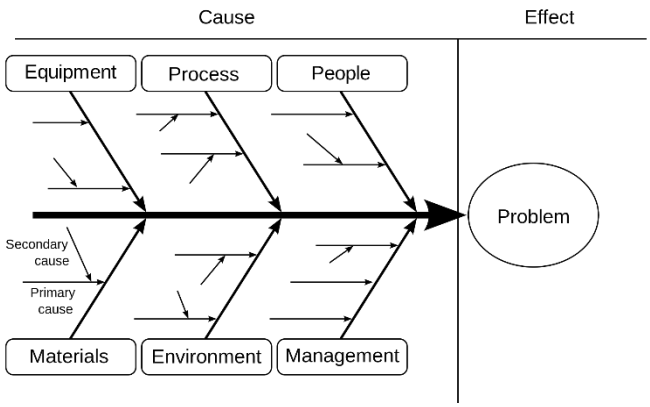
4.2 Data Error Cause Analysis

Using the old analogy that you would be wasting your time cleaning the water in a swimming pool if you had not cured the source of the dirt, similarly for data, it is important to identify root causes of data errors so that these can be resolved before attempting to correct the data. The two core activities in this process are:

- Identifying and correcting root causes – a challenge here is to develop suitable criteria for which data errors to assess in order to avoid getting bogged down in analysis. Once the root cause(s) have been identified, they should be corrected
- Other data stores and systems should be assessed to determine if these root causes could also exist there, in which case they should also be corrected. This activity should help prevent repeat problems.

This process has links to a number of other data quality processes:

- Depending on the error causes identified, there may be a requirement to adjust data stewardship and data flows
- Outputs of the error cause analysis should be an input to data error correction
- Outputs of error cause analysis can be inputs to be the data quality criteria and data designs in order to further reduce the risk of reoccurrence



4.3 Data Error Correction

Data error can be identified from a number of sources, not least the data quality monitoring and data error cause analysis processes. The process involves:

- Agreeing who, how and when data errors will be corrected
- Ensuring that corrections are shared with other parties and data stores as required

For both the above activities, it is important to ensure that staff have the required authorisation to make the corrections.

5 ISO 8000:150-Organisational approaches to Data Quality Management

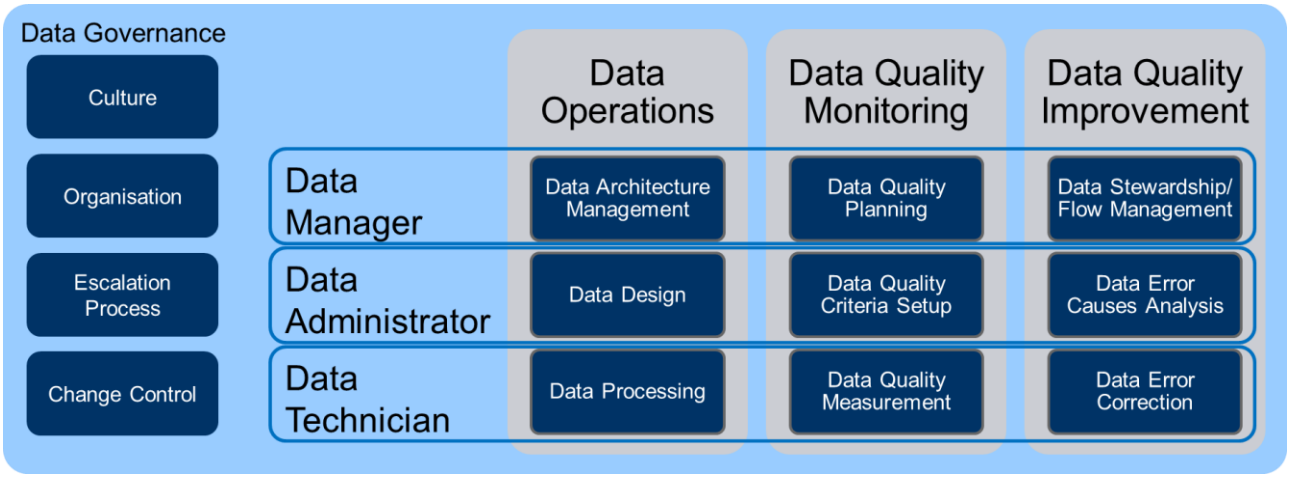
The framework defined in this standard is split into three horizontal 'Roles' with their core responsibilities broadly listed as:

- Data Manager – Responsible for the organisational factors for data quality management and the overall running of the framework. Includes delivery of the three high level processes of Data Architecture Management, Data Quality Planning and Data Stewardship/Data Flow Management
- Data Administrator – Responsible for controlling and coordinating the work of data technicians and aligning to the direction set by the Data Manager. Includes delivery of the three mid-level processes of Data Design, Data Quality Criteria Setup and Data Error Cause Analysis
- Data Technician – In general the data technician is responsible for the actual data changes that are undertaken, specifically as part of the low level processes of Data Processing, Data Quality Measurement and Data Error Correction. Activities will be undertaken under the guidance of the Data Administrator

These roles, as described, are probably one of the areas where I believe that this standard needs most adaptation to 'fit' an organisation. Organisations come in many sizes and structures, therefore, the three roles as described probably only fit a minority of organisations.

Probably a better way to think of these 'Roles' would be to view them as the Strategic, Tactical and Operational activities that are required to develop and deliver your data quality management framework. You then need to determine the individual accountabilities and responsibilities as they apply to your organisation and ensure that staff understand what is expected of them.

6 ISO 8000:150 as a framework for data governance



So far in this White Paper we have reviewed and commented on the general principles and approaches detailed in the standard. This section will be a little different in that it will consider how the standard can form the heart of a coherent approach to data governance, as illustrated in the diagram above.

Some of the things that should be considered for inclusion in this wider data governance framework are as follows:

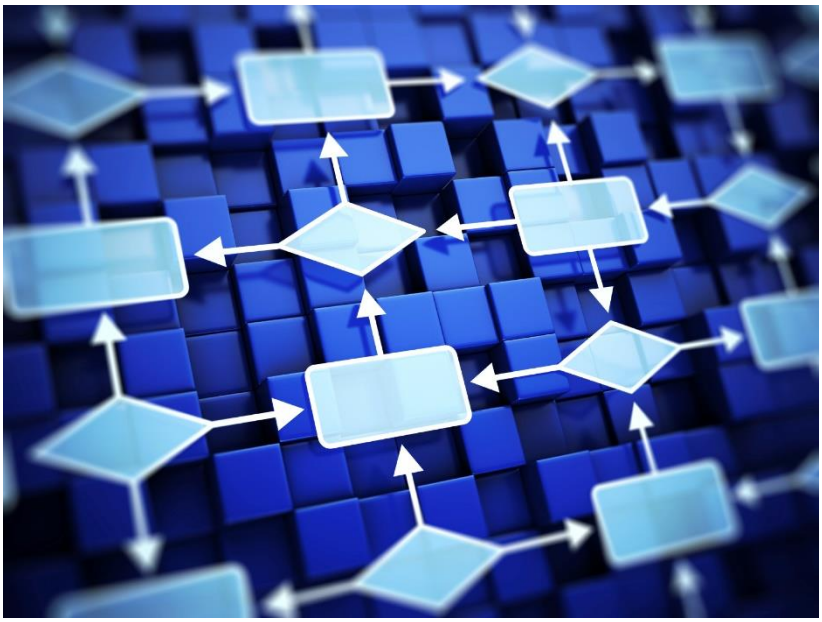
- **Culture** - Do staff, managers and executives understand the importance of data? If data quality issues are identified, what is the typical response? Is good data treated in similar ways to good health and safety approaches?
- **Organisation** – There needs to be clarity on reporting lines between different groups and bodies in an organisation to ensure that nothing 'falls between the cracks' of different areas of influence. This will include ensuring that groups and bodies have a focus on a particular business process or area and not solely on a single software tool or system
- **Executive oversight** – For data to be effectively managed by an organisation, it is vital that there is sufficient senior level oversight of data and data related activities. This can help to ensure that issues are addressed at the appropriate level and that sufficient resources and priority can be applied to data related activities
- **Change management** – Organisations do not exist in a 'steady state' condition, change is on-going and continuous, therefore, data governance and data quality management activities also need to adapt to reflect this change. It is important that sufficient change management and approval is in place to ensure that a 'good idea' does not have long term adverse consequences

- **Escalation** – Problems and issues will arise, so it is important that there is a clear escalation process to allow people to raise issues and ensure they get addressed at the correct level. For any issue there are only three valid outcomes:
  - Resolve the issue, if you have budget and authority to do so
  - Tolerate the issue, if the magnitude does not warrant action
  - Escalate the issue to the next level in the governance hierarchy only if you the resources/authority at the current level cannot resolve the issue and it is also believed that the issue cannot be tolerated at that level

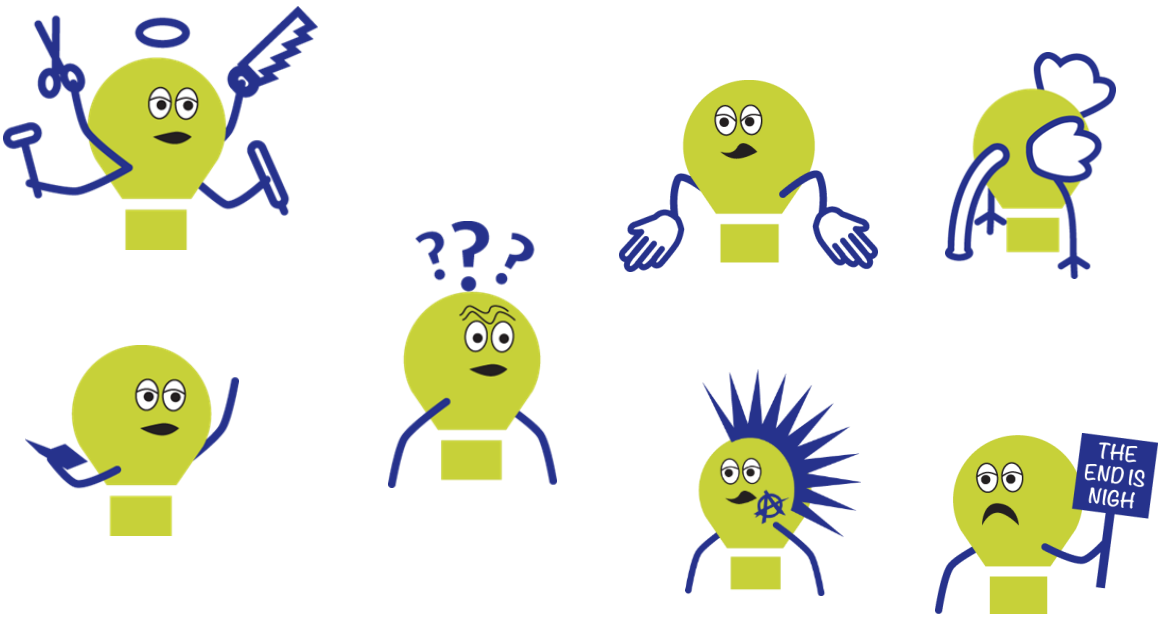
7 Implementation considerations

There are a number of considerations when looking at implementing improved approaches to data governance:

- **Evolution** – Rather than spending a lot of time and resources designing the perfect data governance approach (after a long development time) I am a firm believer in 'light touch data governance'. Link up and strengthen the activities you already have, ensure that there is some level of executive oversight and change management etc. This approach will start to give immediate benefits and can then be adapted and evolved over time.
- **Deliver early benefits** – There may be a number of sceptics in your organisation who may be hoping that they can continue the poor practices they have followed for many years. In order to win over these sceptics, it is important that you identify early benefits delivered by data governance. These early benefits could include the avoidance of data cleansing work, prevention of decisions based on incorrect data or identification of new business opportunities to exploit.
- **Build into existing processes** – Data does not exist in isolation within any organisation – it is an input, output and enabler for most business activities. Since there may already be existing processes, governance groups and dashboards that focus on a specific process area or relate to a particular software tool, it makes more sense to add data in to these processes rather than developing a whole new set of processes and meetings. For staff concerned, this will be seen as a far less intrusive change than a whole new set of data related meetings and processes.



- **Reinforce benefits to other processes** – The more compliant that the culture of an organisation is, the more likely that the approaches to data quality management will support and reinforce approaches to other key areas, such as – Health and Safety, Quality, Risk Management and Performance. These areas are mutually reinforcing – improvements to one area will help secure improvements in other areas.
- **Business As Usual** – Whilst development and implementation of data governance can be thought of as a 'project', data governance should be thought of as a 'Business As Usual' activity in the same vein as Health and Safety Management, Quality Management etc.
- **Whole organisation involvement** – If all parts of an organisation recognise that they are key participants in the approach to data governance, then it is far more likely to achieve the desired outcomes. However, if data governance is seen to be something that a small number of people deliver whilst everyone else continues to do what they have always done, then it is highly likely to fail. If staff are asking "What are YOU going to do about poor data?" then you are not in a good place, however, if staff have an attitude of "What are WE going to do about poor data and how can you help us?" you will probably meet your objectives
- **Culture** – The behaviours of staff towards data will have a significant impact on the quality of data. Our white paper entitled "The Data Zoo – How user behaviours affect information quality" explores this topic in far more detail.



Julian Schwarzenbach is a Director of DPA and has over 20 years' experience in information management in a variety of industries, particularly utilities and transport. Experience has included the management, development and enhancement of enterprise systems, development of data provision processes, data standards, data governance implementations and business transformation.

Julian is Chair of the BCS Data Management Specialist Group and presents regularly on the subject of information management. He managed a project to develop a set of Asset Information Management Guidelines for the Institute of Asset Management and was involved in the development of a range of BIM related standards.

About Data and Process Advantage

Data and Process Advantage (DPA) are experts in helping organisations delivering greater benefits from their data, systems and processes. DPA provide independent pragmatic advice in order to develop positive, long term client relationships. We have extensive experience across a number of sectors, particularly asset intensive sectors including utilities and transportation.

data & process advantage
Data and Process Advantage Limited
Advantage House
Stowe Court
Stowe Street
Lichfield
Staffordshire
WS13 6AQ
England

Telephone: (+44) 1543 401144
Web site: www.dpadvantage.co.uk
Contact: enquiries@dpadvantage.co.uk