

# Asset Information - Data Quality, Confidence and Assurance

## What is the situation?

Data is the “life blood” of an organisation. As it flows between people, systems, databases, processes and departments; it carries with it the ability to make the organisation smarter and more effective. High-performing organisations pay close attention to the data asset, not as an afterthought but as a core part of their business needs. Data is essential to making well-informed decisions that guide and measure realisation of organisational strategy.

Our data is used for multiple activities. For example: by the ORR for official reports, by operations staff to help plan safe work, by the British Transport Police to access the railway and by project teams to plan and deliver their works. However, there is currently a lack of confidence and accuracy in this data meaning that business cases can become unfeasible due to validation requirements. Lack of confidence drives behaviour to create local and competing data sets. This adds costs and prevents the creation of a business-wide view.

## Analysis of causes

STE	1 Data Architecture	Data Quality Planning	2 Data Stewardship & Flow Management	Internal / External Review
AI	Data Design	Data Quality Criteria Setup	Error Cause Analysis	3 Data Confidence Assessment
Routes	Data Processing	4 Data Quality Measurement	Data Error Correction	Management Self Assurance
	No agreed architecture framework - no optimum way found to manage logical data model, no agreed reference architecture at reference level. No optimum way to integrate logical data model outside of the asset domain. Unable to apply 4NF to normalise data and reduce redundancy.	Lack of tools to integrate and profile data across multiple systems. Lack ability to measure conformity with business rules. Inability to forecast data accuracy.	Unable to link business process to organisational accountabilities and data to ensure appropriate people are responsible for data. We lack an integration architecture that is the sum of our application architecture and data architecture to properly enact configuration changes. Lack an ability to measure condition and performance of data movement.	Unable to base assurance on leading indicators. Unable to measure data accuracy. Currently approach uneconomically viable limiting its effectiveness.

## Priority problems

### Specific priority problems

- Data architecture.
- Data quality measurement.
- Data stewardship and flow management.
- Data confidence assessment.

### Related goals

- Create Logical Architecture Framework.
- Requirement for tooling to predict and prevent data quality issues, to improve the accuracy and cost of data that aims to reduce the level of risk in decision making.
- Broadening the scope of linear asset support tooling to integrate a wide range of linear data, to enable decisions that balance the trade-off between performance, cost and risk (TQM).

## Scope

How an organisation uses and manages data is just as important as the mechanisms used to bring it into the business. Having the right data of appropriate quality enables the organisation to perform processes well and to determine which processes have the greatest impact. These fundamental objectives control data by transforming it into useful information. The highest performing organisations ensure that their data assets are accessible to the processes and individuals who need it, are of sufficient quality and timeliness, and are protected against misuse and abuse. This is where we must evolve to.

Successfully controlling data and information assets does not happen by itself. It requires proactive data management by applying specific disciplines, policies, and competencies throughout the life of the data.

Effective data management through all of the data life-cycle phases is the foundation of reliable information. Data may have different uses at different times and requires different management handling in the life-cycle phases. Data without context has no value. Data that consumers never use is worthless too. The value of data is in the information it contains and uses.

## Value of data

How can we assign a meaningful value to data and data quality that enables us to determine a business case to drive improvements and enhancements? How do we link changes in data quality to company risk measures/liabilities? How would we financially value the data as an item on the balance sheet? What is the realised and potential value of our data? How could we insure it against loss?

## Data as an asset

We aim to treat data as an asset in its own right. How do we understand the value of our data so we can invest in it appropriately? How do we calculate the whole life cost of data so we can continue to operate at lowest whole life cost? How do we prevent errors before they occur? How do we move from a “find and fix” to a “predict and prevent” approach to data quality? How do we create decision support capability that helps us describe the failure, diagnose its cause, prognoses when risks will materialise, prescribe appropriate action and ultimate automate the response? How do we link data quality to the softer more qualitative aspects of an organisation? E.g. How does a culture affect data quality and decision making? How do we assure ourselves our data is accurate without resorting to expensive and slow manual validation exercises?

## Data leadership & information engineers

We need to professionalise our data management discipline to offer people a career in data or a platform to move across the organisational disciplines. What do our Information Engineers of the future look like? What skills/behaviours should our people have so they can use/manage the data? How do we attract the best data management talent? What do we teach at school today to prepare the youth to be data leaders of tomorrow? How do we instil the importance of data in today's workforce?

## Linking data requirements to business process

Data offers no value unless it's linked to a purpose. How do we link data into business processes? How do we link data to the risk, objectives and outputs of Network Rail? How do we link data to the social value of the organisation that is “beyond the balance sheet”.

## How to deal with data uncertainty

The goal is not perfect data, it is understanding how imperfect the data is so that risk can be factored into decision-making. Waiting for our feet data before making a decision will lead to stagnation. The challenge is how do we make the best use of imperfect data sets, whilst still being confident in the outcome? Can we build in learning/experience to improve predictions and reduce risk?