

NPL Report Number MS 63

Data Quality Good Practice Guide

João Gregório and Paul Duncan

with contributions from

Moulham Alsuleman, Michael Chrubasik, Bartłomiej Cieszynski, Vivek Desai, Jean-Laurent Hippolyte, Padmini Krishnadas, Keith Lines, Russell Miller, Nina Perić, and Harvey Whelan

March 2026



National Physical Laboratory (NPL)

Data Quality Good Practice Guide

by João Gregório and Paul Duncan

with contributions from (*alphabetically*):

**Moulham Alsuleman, Michael Chrubasik, Bartłomiej Cieszynski, Vivek Desai,
Jean-Laurent Hippolyte, Padmini Krishnadas, Keith Lines, Russell Miller, Nina Perić,
and Harvey Whelan**

**We are the UK's National Metrology Institute (NMI),
a world-leading centre of excellence that provides cutting-edge
measurement science, engineering and technology that
underpins prosperity and quality of life in the UK.**

© NPL Management Limited, 2026

ISSN 1754-2960

<https://doi.org/10.47120/npl.MS63>

National Physical Laboratory
Hampton Road, Teddington, Middlesex, TW11 0LW

This work was funded by the UK Government's Department for Science, Innovation & Technology through the UK's National Measurement System programmes.

Extracts from this report may be reproduced provided the source is acknowledged and the extract is not taken out of context.

Executive Summary

Data quality underpins effective decision-making across multiple domains of application. As organisations increasingly rely on digital systems, from analytical pipelines and digital twins to regulatory reporting platforms, the consequences of poor-quality data become more pronounced. This guide provides good practice guidance for understanding and assessing data quality, grounded in established international standards and informed by practical applications across diverse domains.

The guide adopts the ISO 25012 data quality model as its foundation, which defines 15 data quality dimensions organised along a spectrum from inherent to system-dependent. This framework is extended with four additional dimensions (governance, usefulness, quantity, and semantics) identified through recent research as increasingly relevant to modern data applications. The resulting set of 19 dimensions provides comprehensive coverage of the contemporary data quality landscape while maintaining the structured classification that supports its systematic assessment.

Beyond dimension definitions, the guide surveys existing data quality frameworks, comparing generalisable standards such as ISO 25012 and ISO 8000 with domain-specific frameworks such as ALCOA+ for life sciences and BCBS 239 for financial services. A comparative analysis shows that core dimensions, such as accuracy, completeness, consistency, and currentness, are steadily represented across frameworks, while dimensions such as semantics, quantity, and governance are less represented. This analysis provides guidance on framework selection, helping organisations identify approaches suited to their regulatory contexts and operational requirements.

A central theme of this guide is that data quality is inherently contextual. The 19-dimension structure offers a common vocabulary for discussing data quality, but the practical interpretation of each dimension varies depending on the application environment. Six case studies drawn from manufacturing, healthcare, maritime, and environmental monitoring domains are presented in this guide to illustrate this principle. The data quality dimension of traceability, for instance, takes on different meanings across these case studies: in the assessment of simulation credibility it is related to the provenance of documentation; in supply chain integration it relates to the integrity of transferred files; and

in environmental monitoring it is attached to regulatory compliance. Recognising these contextual dependencies is essential for implementing data quality practices that are aligned with established frameworks and tailored to specific operational needs.

Future developments will explore the formalisation of dimension operationalisation through structured representations of how dimensions are interpreted across different domains of application. The long-term vision is a living framework, based on semantic web standards, that captures validated interpretations contributed by practitioners, creating a shared resource that bridges abstract definitions and concrete assessment practices.

This guide is intended for data practitioners, quality managers, researchers, and anyone responsible for ensuring that data are fit for purpose. It provides both conceptual groundwork and practical examples that support the design and implementation of effective data quality programmes.

Contents

Executive Summary	3
Contents	5
Purpose and Scope of this Guide	6
An Introduction to Data Quality	8
Uncertainty in Measurement Data	10
Data Quality Dimensions	13
Data Quality Frameworks	25
Case Studies	32
Conclusions	47
Future Work	49
Acknowledgements	51
References	52

Purpose and Scope of this Guide

This Good Practice Guide provides comprehensive guidance for assessing and managing data quality across diverse applications and domains. It is intended for data professionals, quality assurance personnel, managers, and anyone responsible for ensuring data quality in their organisations.

Trustworthy data-driven decision-making, regulatory compliance, and operational effectiveness all depend on systematic approaches to data quality. Although the guidance presented in this document was developed with NPL's role as the UK's National Measurement Institute in mind, the principles and approaches presented are applicable across various industries and organisational contexts. The framework builds upon established international standards, particularly ISO 25012 due to its generalisable nature, while incorporating additional dimensions and contemporary best practices relevant to modern digital environments.

As businesses across all sectors evolve towards data-driven operations, the volume, velocity, and complexity of data continue to grow exponentially. The increasing dependence on digital systems, particularly artificial intelligence (AI) and machine learning (ML) for informed decision-making, makes high-quality data essential not only for operational efficiency but also for increasing trust in these systems.

Systematic data quality management offers distinct benefits across different stakeholder groups. Data creators gain enhanced confidence in their outputs, reduced risk of errors and rework, and improved efficiency in collection and processing workflows. Data users benefit from greater trust in data-driven decisions through transparent quality indicators, improved ability to assess fitness-for-purpose, and enhanced interoperability supporting FAIR data principles. For organisations more broadly, comprehensive data quality practices strengthen reputation and credibility, improve operational efficiency, and ensure compliance with international standards and regulatory requirements.

Against this backdrop, the primary objectives of this guide are to present a standardised framework of data quality dimensions organised into distinct categories, to provide practical guidance for implementing data quality assessment programmes, to demonstrate

their applications through real-world examples, and to facilitate effective communication about data quality across different domains and stakeholders.

Coverage spans the complete data lifecycle, from collection and processing through analysis, storage, and dissemination. The framework applies to all forms of data, including structured and unstructured datasets, measurement and experimental data, business and operational records, research and scientific data, and digital assets and metadata.

Technical scope extends to both traditional and emerging technologies, including cloud-based systems, artificial intelligence applications, and Internet of Things deployments.

Rather than prescribing specific technical implementations or software solutions, this guide focuses on assessment and management practices. Organisations should adapt the framework to their specific operational contexts, regulatory requirements, and stakeholder needs while maintaining alignment with the fundamental principles presented.

An Introduction to Data Quality

The increasing dependence on data for informing decision-making processes brings with it a growing need to ensure that data are fit for purpose [1]. This reliance spans virtually every sector, from pharmaceutical manufacturing and healthcare to engineering and education [2]. As organisations integrate increasingly complex digital systems into their operations, the consequences of poor data quality become more pronounced [3]. For instance, modern analytical pipelines, digital twins, and predictive maintenance systems rely heavily on accurate, timely, and well-structured data to function as intended [4] [5].

Even contemporary AI-enabled tools – although far from the only example – highlight the same underlying principle: the performance and reliability of these systems are fundamentally constrained by the quality of the data they consume [6]. This makes it an opportune moment to emphasise the broader point that robust data governance and quality management are essential, regardless of whether the system is a machine learning pipeline, a process control system, or a regulatory decision support platform [7].

Addressing these challenges requires a clear understanding of what data quality means and how it can be evaluated. This guide adopts the definition established by ISO 25012, where data quality represents the degree to which data possess the characteristics necessary to satisfy stated and implied needs when used under specified conditions [8]. It is defined by a set of dimensions that can be qualitatively or quantitatively assessed to describe the quality of datasets and other data structures. Common dimensions include accuracy, completeness, traceability, and timeliness, though the terminology used to describe them is complex, leading to ongoing standardisation efforts [9]. In measurement science, data quality encompasses not only the accuracy and precision of numerical values but also the completeness of metadata, the consistency of formats, and the reliability of data provenance [10] [11].

Data quality is also inherently contextual, meaning that data suitable for one application may be inadequate for another [12]. This context-dependency is particularly important in metrology, where the same measurement data might be used for routine quality control, regulatory compliance, scientific research, and commercial transactions each with different quality requirements. Despite efforts to establish standardised metrics and assessment

procedures, different industries have unique data needs, and inconsistencies in defining and evaluating quality persist across sectors. This lack of standardisation has resulted in considerable diversity in existing terminology, which creates communication challenges when addressing data quality across domains. This guide draws upon recent work that maps the diverse terminology used across sectors into a unified framework, enabling data practitioners to communicate about data quality using a common vocabulary [1].

Implementing such a unified understanding in practice typically relies on data quality frameworks. These are structured methodologies used to assess, manage, and improve the quality of data. They can be built upon existing standards and regulations or specifically designed for more tailored applications [2]. These frameworks are essential for organisations to manage their data and demonstrate evidence of good data quality practices when communicating with stakeholders. Frameworks range from generalised approaches applicable across multiple domains and use cases to specialised frameworks addressing specific regulatory and operational requirements within sectors [7]. While specialised frameworks typically cover fewer overall dimensions than generalisable frameworks, they often include dimensions absent from broader frameworks, providing targeted insights crucial for their specific applications [13]. While these frameworks provide valuable structure for assessing data quality broadly, measurement data possess unique characteristics that warrant specific consideration.

Uncertainty in Measurement Data

Assessing and quantifying the quality of data requires an estimation of the "confidence" in the process that generated that data. In metrology, the science of measurement, this practice is already well established and provides a useful foundation for data quality assessment. Metrology follows mature principles, including the Guide to the Expression of Uncertainty in Measurement (GUM) and the International Vocabulary of Metrology (VIM), which offer structured frameworks for understanding and expressing measurement quality [10] [11]. These distinctive characteristics mean that data quality assessment in measurement contexts must consider metrological principles alongside standard data quality dimensions.

All measurement results include associated uncertainty information that quantifies the reliability of the reported value, and this uncertainty must be clearly documented and propagated through data processing workflows. Figure 1 illustrates the relationship between the measured value, the measurand (the "true" value being measured), the error, and the uncertainty. Measurement results must also demonstrate traceability to recognised

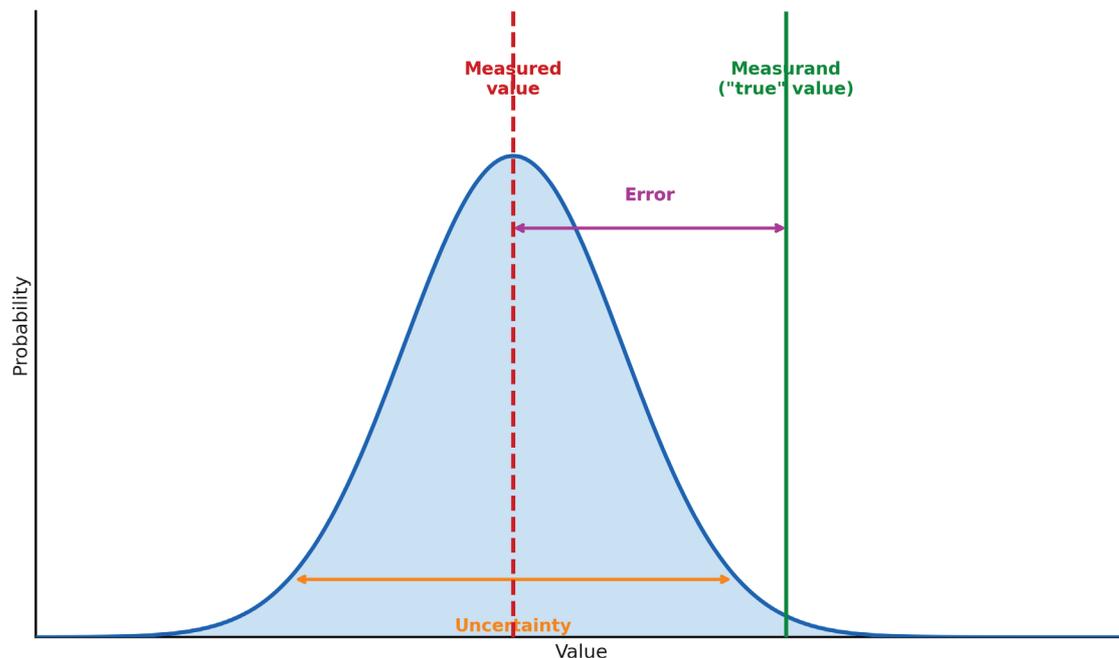


Figure 1: Distinction between error and uncertainty. Error is the difference between the measured value and the measurand, whereas uncertainty quantifies the confidence in the measurement result.

national measurement standards through an unbroken chain of calibrations, and this traceability information forms essential metadata for data quality assessment.

Measurement results are also critically influenced by environmental conditions such as temperature, humidity, and pressure, making complete documentation of these conditions necessary for correct interpretation. The calibration status of equipment introduces an additional temporal aspect, in the form of a calibration timestamp, that can affect measurement quality. Figure 2 shows how these metadata elements move through a typical measurement system and highlights the consequences of incomplete documentation.

Understanding how uncertainty propagates through a measurement system provides insight into how confidence accumulates, or diminishes, across connected processes. In metrology, every measured quantity is treated as having an associated uncertainty, and mathematical models describe how these uncertainties combine to influence outputs. Within the GUM, the model function, sensitivity coefficients and covariance structure make explicit how each input's distribution influences the output, providing a structured basis for assessing the "fitness for purpose" of data quality dimensions. In this way, uncertainty information functions as structured metadata, comparable to other data quality dimensions such as accuracy and completeness. Uncertainty budgets therefore encompass not only expanded uncertainties but also the underlying standard uncertainties, correlation assumptions and model specifications in machine-readable form.

As measurement systems increasingly become part of automated workflows, digital twins, and AI processes, metrological principles can be embedded within these systems to ensure that uncertainty is not only calculated but also propagated through complex digital structures. Concepts such as uncertainty quantification and uncertainty propagation therefore provide a foundation for expressing confidence across data quality dimensions, extending the rigour of metrology to modern, data-centric, and interconnected measurement environments.

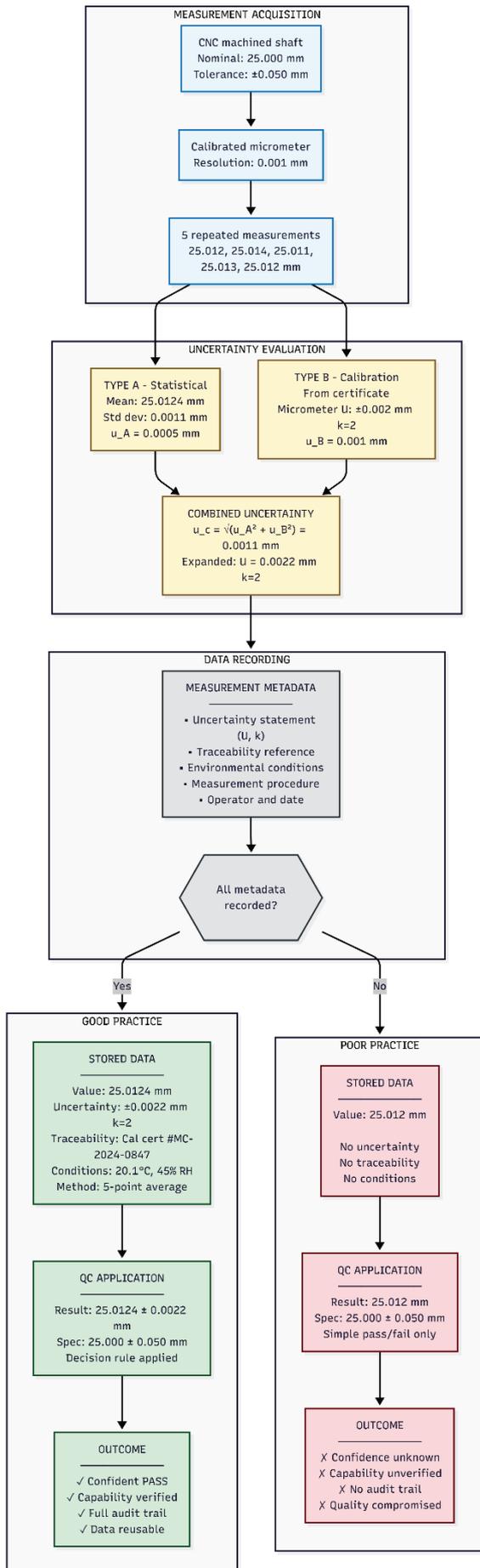


Figure 2: Measurement data lifecycle showing the impact of metadata preservation on data quality in a measurement context. A shaft diameter measurement shows how uncertainty is evaluated from Type A (statistical analysis of repeated measurements) and Type B (calibration certificate) components. At the data recording stage, measurement metadata must be captured: the uncertainty statement with coverage factor, traceability reference, environmental conditions, measurement procedure, and operator details. The good practice path (left) preserves this metadata, enabling confident conformance decisions with applied decision rules and maintaining a complete audit trail. The poor practice path (right) records only the measured value, rendering the data unable to support reliable quality decisions or demonstrate measurement validity.

Data Quality Dimensions

Data quality dimensions are specific, quantifiable characteristics used to assess and describe the quality of data within a given context. Each dimension represents a distinct aspect of data quality that can be evaluated independently, yet collectively they provide a view of overall data fitness for purpose. The 19 dimensions presented in this guide are organised into three categories – Inherent, contextual, and system-dependent – based on their nature and their relationship to data usage contexts [1]. This categorical structure reflects how quality characteristics manifest differently depending on whether they arise from the data itself, its context of use, or the systems that manage it, helping practitioners identify which dimensions are most relevant to their specific assessment needs. Figure 3 shows this hierarchical structure, with the outer ring displaying the diverse terminology used across different frameworks and domains that map to each standardised dimension. These dimensions have also been formalised into a structured taxonomy, available as an



Figure 3: Data quality dimensions classified into three categories based on their characteristics: inherent dimensions (red); contextual dimensions (blue); and system-dependent dimensions (green).

interactive resource¹ where practitioners can explore definitions, examples, and relationships between categories. This section provides detailed definitions and guidance for each dimension, enabling practitioners to understand, assess, and improve specific aspects of their data quality systematically.

Inherent dimensions

Inherent data quality represents the degree to which the data's quality characteristics have the intrinsic potential to meet stated and implied needs when used under specified conditions. Errors in the data can propagate through workflows, causing significant problems when data are used for reporting, analytics, and decision-making. Addressing inherent data quality requires robust validation against established rules, appropriate governance oversight, and thorough documentation. By examining dimensions such as accuracy, completeness, consistency, credibility, and currentness, organisations can verify that their data integrity measures enable data to meet quality expectations.

Accuracy

Accuracy is the degree to which data have attributes that correctly represent the true value of the intended attribute of a concept or event in a specific context of use [1] [8]. In measurement science, accuracy reflects how closely a measured or recorded value corresponds to the true or reference value. This dimension is fundamental to data quality, as inaccurate data can lead to flawed analyses, incorrect conclusions, and poor decisions.

Example(s) and possible interpretation(s):

Accuracy: A temperature sensor used in an environmental chamber can be assessed for accuracy by comparing its readings against a calibrated reference thermometer traceable to national standards. If the sensor reads 23.2 °C while the reference thermometer (with known uncertainty) reads 23.0 °C ± 0.1 °C, the accuracy of the sensor can be quantified.

¹ Interactive Data Quality Taxonomy: <https://dq-taxonomy-7aa56b.gitlab-docs.npl.co.uk/>. This is a live resource under active development; if inaccessible, please contact the corresponding author.

Completeness

Completeness is the degree to which subject data associated with an entity have values for all expected attributes and related entity instances in a specific context of use [1] [8]. This dimension addresses whether all required data elements are present and populated. Incomplete data can result from collection failures, system errors, or deliberate omissions. In measurement contexts, completeness ensures that all necessary parameters, metadata, and supporting information are captured alongside primary measurements.

Consistency

Consistency is the degree to which data have attributes that are free from contradiction and are coherent with other data in a specific context of use [1] [8]. It can apply both amongst data regarding one entity and across similar data for comparable entities. Consistent data maintain logical relationships and follow established rules across datasets, time periods, and systems. In measurement applications, consistency ensures that related measurements, units, and formats align correctly throughout data lifecycles.

Example(s) and possible interpretation(s):

Completeness: A materials testing database can be checked for completeness by identifying records with missing mandatory fields. For example, tensile strength measurements that lack associated specimen dimensions, test temperature, or humidity conditions. The completeness can be quantified as the percentage of records with all required fields populated.

Consistency: In a chemical analysis dataset, the sum of constituent percentages for a compound must equal 100 %. For example, a sample listing constituents at 45 %, 30 %, and 25 % is internally consistent because the values add up exactly to 100 %. In contrast, a record showing constituents that total 110 % indicates a consistency error, suggesting problems such as incorrect measurements, rounding mistakes, or duplicated entries.

Credibility

Credibility is the degree to which data have attributes that are regarded as true and believable by users in a specific context of use [1] [8]. Credibility includes the concept of authenticity, encompassing the truthfulness of origins, attributions, and commitments. This dimension reflects user trust in data sources and processes. In metrology, credibility is closely linked to traceability to national standards, calibration records, and the reputation of measurement facilities and personnel.

Currentness

Currentness is the degree to which data have attributes that are of the right age in a specific context of use [1] [8]. Also referred to as timeliness or temporality, this dimension assesses whether data are sufficiently up-to-date for their intended purpose. The acceptable age of data varies significantly by application. For example, real-time monitoring requires immediate data, while historical analyses may use decades-old records. In measurement contexts, currentness includes consideration of calibration validity periods and measurement timestamps.

Example(s) and possible interpretation(s):

Credibility: Measurement data from an accredited calibration laboratory carries higher credibility than data from an unknown source. Credibility can be assessed by evaluating laboratory accreditation status, participation in proficiency testing programmes, and measurement and calibration capability listings in databases.

Currentness: A mass balance used for high-precision measurements requires annual calibration. Its currentness can be assessed by checking the calibration due date: if the current date is 15 March 2025 and the last calibration was performed on 20 March 2024 with a 12-month validity period, the calibration remains current until 20 March 2025. Measurements from equipment whose calibration has expired represents a failure that compromises traceability.

Contextual dimensions

Contextual data quality emphasises evaluating the quality of data within the specific context in which it is used [1] [8]. This category recognises that whether data are fit for purpose depends on the application environment, user requirements, and operational constraints. For example, in a measurement laboratory setting, calibration data must be readily accessible to technical staff for timely decision-making, yet rigorous controls must prevent unauthorised modification, illustrating the interplay between accessibility and compliance. Considering contextual dimensions allows a more complete assessment of overall fitness for purpose, highlighting how key quality issues relate to the environment in which data are used.

Accessibility

Accessibility is the degree to which data can be accessed in a specific context of use, particularly by people who need supporting technology or special configuration because of some disability or additional needs [1] [8]. This dimension encompasses both technical accessibility (can systems retrieve the data?) and human accessibility (can users with varying abilities interact with the data?). In measurement science, accessibility ensures that data are available to authorised personnel through appropriate interfaces, formats, and tools while maintaining necessary security controls.

Example(s) and possible interpretation(s):

Accessibility: A research dataset containing spectroscopy measurements is stored as proprietary binary files that require specialised viewing software costing £5,000 per licence. Accessibility can be improved by converting data to open formats such as CSV or HDF5, albeit at the cost of losing accuracy. Additionally, providing comprehensive metadata in human-readable formats (JSON, XML) and ensuring web interfaces meet accessibility standards enables users with assistive technologies to discover and retrieve data effectively.

Compliance

Compliance is the degree to which data have attributes that adhere to standards, conventions, or regulations in force and have similar rules relating to data quality in a specific context of use [1] [8]. This dimension addresses conformance to legal requirements, industry standards, organisational policies, and technical specifications. For measurement data, compliance may involve adherence to ISO standards, regulatory requirements (such as good laboratory practices), data format specifications, and documented quality management procedures.

Confidentiality

Confidentiality is the degree to which data have attributes that ensure that they are only accessible and interpretable by authorised users in a specific context of use [1] [8]. This dimension protects sensitive information from unauthorised access or disclosure. In measurement environments, confidentiality concerns may include commercially sensitive calibration data, personal information in medical measurements, or security-critical measurement results

Example(s) and possible interpretation(s):

Compliance: A pharmaceutical laboratory conducting stability testing must demonstrate compliance regulations. This assessment involves verifying that: measurement data include all mandatory metadata fields (electronic signatures, timestamps, audit trails); environmental conditions are recorded at specified intervals; all instruments have current calibration certificates; and data formats conform to regulatory submission requirements.

Confidentiality: Medical device testing data containing patient information must undergo statistical disclosure control before publication. Assessment involves verifying that: direct identifiers are removed; quasi-identifiers are generalised; and k-anonymity analysis confirms no combination of attributes can identify fewer than $k=5$ individuals.

Efficiency

Efficiency is the degree to which data have attributes that can be processed and provide the expected levels of performance using the appropriate amounts and types of resources in a specific context of use [1] [8]. This dimension considers data size, format optimisation, processing speed, and resource consumption. Efficient data structures enable timely analysis without excessive computational or storage costs. In measurement applications, efficiency considerations include file formats, compression methods, and database indexing strategies that support rapid retrieval and analysis.

Governance

Governance is the degree to which data have attributes that adhere to the formalised frameworks of authority and accountability that support harmonised data activities across an organisation [1] [8]. This dimension addresses the policies, procedures, roles, and responsibilities that guide data management throughout its lifecycle. Effective data governance ensures consistent approaches to data quality, security, retention, and usage. In measurement laboratories, governance frameworks define who can create, modify, approve, and delete measurement records, as well as how data quality is monitored.

Example(s) and possible interpretation(s):

Efficiency: A measurement database containing 10 million calibration records experiences slow query performance when retrieving data for specific instruments over date ranges. Efficiency can be assessed by measuring query execution time and database size. Implementing appropriate indexing on specific fields and applying data compression can reduce query time and storage.

Governance: An organisation preparing for an audit requires audit trails of data access and changes. Governance assessment involves evaluating whether auditability mechanisms are in place to track timestamps, data field changes, and the source user or system making edits. This transparency enables verification of compliance, upholding accountability of data handling in the data lifecycle.

Precision

Precision is the degree to which data have attributes that are exact or that provide discrimination in a specific context of use [1] [8]. In measurement science, precision describes the closeness of agreement between independent measurement results obtained under specified conditions. This dimension encompasses concepts such as granularity, resolution, and the level of detail captured in data. High precision enables fine distinctions but may not guarantee accuracy.

Traceability

Traceability is the degree to which data have attributes that provide an audit trail of access to the data and of any changes made to the data in a specific context of use [1] [8]. This dimension encompasses provenance (origin and history), lineage (transformation and processing steps), and auditability (who did what, when). In metrology, traceability extends to the metrological concept of an unbroken chain of calibrations linking measurements to national or international standards. Comprehensive traceability supports data verification, error investigation, and regulatory compliance.

Example(s) and possible interpretation(s):

Precision: Precision can be assessed through repeatability measurements. Ten consecutive mass measurements of a 100 g reference weight yield a standard deviation of 0.001 g, indicating that measurements cluster tightly. This level of precision is fit for purpose for applications requiring discrimination to 0.01 g, demonstrating adequate repeatability for the intended use.

Traceability: A laboratory force gauge calibration must demonstrate metrological traceability to national standards. Assessment involves verifying an unbroken chain: the force gauge was calibrated against a reference load cell, which holds a valid calibration certificate traceable to the UK national standard. Each link includes calibration certificates and uncertainty budgets.

Understandability

Understandability is the degree to which data have attributes that enable them to be read and interpreted by users, as well as are expressed in appropriate languages, symbols, and units in a specific context of use [1] [8]. This dimension addresses clarity, documentation, labelling, and the use of recognised conventions. Well-documented metadata, clear variable names, and appropriate visualisations enhance understandability. In measurement contexts, understandability includes unit specification, uncertainty statements, and enough contextual information for users to interpret measurement results.

Usefulness

Usefulness is the degree to which data attributes meet the specific requirements of users or applications [1] [8]. This includes the data's adaptability across various contexts, recognising their potential for diverse applicability due to aspects such as reusability and interoperability. Useful data are relevant, fit for purpose, and provide value to stakeholders. This dimension considers whether data answer the questions being asked and support intended decisions. In measurement science, usefulness encompasses whether measurements have appropriate ranges and uncertainties for example.

Example(s) and possible interpretation(s):

Understandability: A dataset of thermal conductivity measurements includes values but lacks sufficient contextual information. Units are not specified, measurement conditions are absent, and variable names are cryptic. Understandability requires clear units, metadata, and variable names.

Usefulness: A researcher requires particle size distribution data for validating a fluid dynamics model, needing measurements with uncertainties below 5% covering a 1–100 μm diameter range. Usefulness assessment considers whether measurement range, uncertainty, and accompanying metadata collectively meet the specific requirements of the intended application.

System-dependent dimensions

System-dependent quality refers to aspects of data quality that are influenced by the technological and organisational systems that collect, store, manage, and provide access to data [1] [8]. These dimensions can vary substantially based on the capabilities and constraints of interconnected data sources, processing pipelines, and storage infrastructure. For measurement organisations, this aspect is particularly relevant as digital transformation introduces increasingly complex data architectures spanning laboratory management systems (LIMS), calibration databases, and external data exchanges. Addressing system-dependent data quality requires a comprehensive understanding of the system landscape, including establishing clear data definitions for consistent semantics, implementing reliable backup and recovery processes, and designing flexible data architectures to support portability across evolving technological environments.

Availability

Availability is the degree to which data have attributes that enable them to be retrieved by authorised users and/or applications in a specific context of use [1] [8]. This dimension captures whether users can access the data they need when they require them. Availability encompasses both the user's ability to find, access, and retrieve data, as well as the adequacy of data available to a given user. In measurement contexts, availability considerations include whether calibration records, measurement results, and associated metadata can be accessed by relevant personnel across different locations and systems.

Example(s) and possible interpretation(s):

Availability: Consider an organisation that has databases distributed across different regions. Data availability may be high within each local system but limited when attempting access from a different region due to network constraints or data sovereignty laws. An availability assessment involves evaluating whether authorised users can retrieve the data they need when required, regardless of their location or the system they are using.

Portability

Portability is the degree to which data have attributes that enable them to be installed, replaced, or moved from one system to another whilst preserving existing quality in a specific context of use [1] [8]. This dimension concerns preserving the utility and meaning of data when moving across storage, software, and hardware. Portable data use open standards and well-documented formats that facilitate long-term preservation and system migrations. For measurement data, portability ensures that records remain usable as technology evolves, supporting requirements for long-term retention.

Quantity

Quantity is the degree to which data have attributes that represent sufficient amount or volume, providing a comprehensive view of the intended attribute of a concept or event in a specific context of use [1] [8]. This dimension assesses whether the amount and coverage of available data is sufficient for its intended application. Too little data may lead to insufficient statistical power or incomplete coverage, whilst excessive data may create storage, processing, and management challenges. In measurement contexts, quantity considerations include sample sizes, measurement frequencies, and the extent of measurement campaigns required to achieve statistically meaningful results.

Example(s) and possible interpretation(s):

Portability: Measurement data stored in a proprietary binary format cannot be readily transferred to a collaborating organisation's systems. Converting to open formats with well-defined schemas improves portability, enabling seamless data exchange whilst minimising the risk of corruption or loss during transition.

Quantity: An environmental monitoring campaign must determine appropriate measurement frequencies and sampling durations. Quantity assessment evaluates whether captured data provide adequate temporal and spatial coverage to draw statistically meaningful conclusions.

Recoverability

Recoverability is the degree to which data have attributes that enable them to maintain and preserve a specified level of operations and quality, even in the event of failure, in a specific context of use [1] [8]. This dimension encompasses backup procedures, disaster recovery capabilities, and data resilience mechanisms that rapidly restore services with minimal data loss. For measurement laboratories, recoverability is critical for maintaining unbroken records required by quality management standards and for preventing loss of irreplaceable data such as historical calibration records.

Semantics

Semantics is the degree to which data accurately and consistently represent the intended meaning, interpretation, and real-world concepts within a specific context of use, ensuring correct semantic understanding by users and applications [1] [8]. Semantic quality becomes particularly important in data integration, knowledge graphs, and machine learning applications where automated systems must interpret data correctly. In measurement science, semantics ensures that quantities, units, and relationships are unambiguously represented, supporting interoperability across measurement infrastructures.

Example(s) and possible interpretation(s):

Recoverability: Historical calibration records spanning decades require safeguards against deterioration over extended retention cycles. Recoverability assessment evaluates measures to protect data integrity against decay, ensuring archived records remain accessible for future traceability requirements.

Semantics: An electronic health record system must ensure that patient data are consistently represented across all sections. If a blood type is recorded as "A+" in one section but "A-" in another, this semantic inconsistency could lead to errors. Semantic assessment verifies that data represent real-world entities without ambiguity and that meaning remains consistent over time.

Data Quality Frameworks

Data quality frameworks are structured methodologies used to assess, manage, and improve the quality of data. They can be directly built upon existing standards and regulations, or they can be specifically designed to address specific applications.

Generalisable frameworks

Total Data Quality Management (TDQM) is a holistic strategy developed at MIT in the 1980s that views data as a commodity and employs methodologies to ensure high quality [14]. TDQM was a pioneering research programme on data quality and has significantly influenced the development of the field. The framework focuses on different dimensions of data quality that correspond to data quality categories, including accuracy, objectivity, believability, reputation, access, security, relevance, value-added, timeliness, completeness, amount of data, interpretability, ease of understanding, concise representation, and consistent representation. The application process comprises four stages through a Define, Measure, Analyse, Improve (DMAI) cycle, shown in Figure 4. The definition phase involves determining the relevant dimensions of data quality for both the organisation and the specific data being considered. The measurement phase assesses the existing condition of data quality, identifying any issues and understanding their effects. The analysis phase investigates the fundamental reasons behind data quality challenges. Finally, the improvement phase executes modifications that address the identified challenges in order to enhance data quality. TDQM covers 11 out of 19

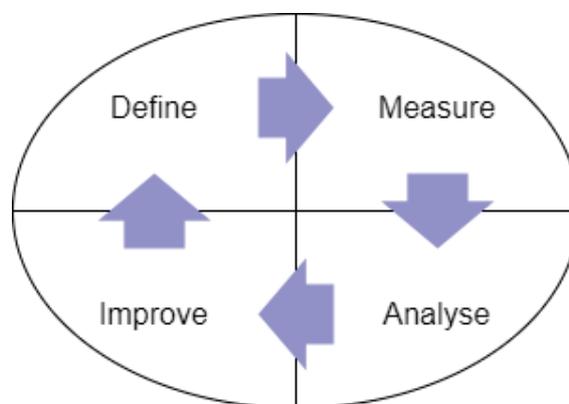


Figure 4: “Define, Measure, Analyse, Improve” cycle as outlined by TDQM for refining adequate data management processes to implement.

dimensions presented previously in this guide and is one of the earliest examples of a structured data quality framework.

ISO 8000 is an international standard that provides a comprehensive approach to data quality management [7]. It is part of the broader ISO 9000 series focused on quality management systems [15]. ISO 8000 follows a "Plan, Do, Check, Act" iterative cycle, shown in Figure 5, that leads to incremental change and improvements in data management processes. This framework can be applied across various industries and data types, providing guidelines for data quality requirements, data governance, and data management. ISO 8000 emphasises the importance of quality data in supporting business processes and decision-making.

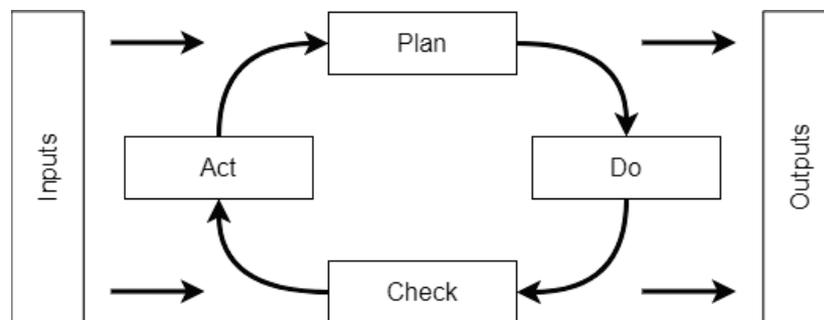


Figure 5: "Plan, Do, Check, Act" cycle, outlined by the ISO 8000 and ISO 9000 series. The cycle follows an iterative process leading to incremental change and continual improvements and can be used to design better data management processes.

ISO 25012 is part of the Software Product Quality Requirements and Evaluation (SQuaRE) series of international standards [8]. This standard establishes a general-purpose data quality model that can be applied to data stored within a structured computer system. ISO 25012 lists 15 basic data quality traits: accuracy, completeness, consistency, credibility, currentness, accessibility, compliance, confidentiality, efficiency, precision, traceability, understandability, availability, portability, and recoverability [8]. These dimensions are classified into a perspective spectrum ranging from inherent to system-dependent. Inherent data quality dimensions represent fundamental, intrinsic properties of data that hold true regardless of context or user requirements. System-dependent dimensions underscore the role of the system in data quality, including aspects such as data availability, portability, and recoverability. The comprehensive approach to ISO 25012 enables organisations to employ a single framework for consistently assessing data quality while maintaining sufficient granularity through its 15 individual dimensions. This standard

forms the foundation for many other frameworks and provides the basis for the extended 19-dimension framework presented in this guide.

Domain-specific frameworks

ALCOA+ is a data integrity framework widely adopted in life sciences and endorsed by the US Food and Drug Administration (FDA) [16]. The letters in the acronym stand for Attributable, Legible, Contemporaneous, Original, and Accurate, as well as Complete, Consistent, Enduring, and Available. ALCOA+ addresses data quality needs specific to pharmaceutical development, clinical trials, and regulatory compliance in healthcare. The framework emphasises traceability, understandability, and governance as particularly relevant to regulated medical environments. While it effectively addresses life science requirements, its applicability to other sectors is limited due to its domain specificity.

BCBS 239 is the Basel Committee on Banking Supervision's standard for risk data aggregation and risk reporting, specifically designed for the financial sector [17]. This framework addresses the need for accurate, complete, and timely risk data to support effective risk management and regulatory reporting in banking institutions. BCBS 239 emphasises dimensions including accuracy, completeness, timeliness, consistency, and governance, with particular attention paid to data distribution and aggregation across complex banking operations. The framework recognises that effective risk management requires high-quality data that can be aggregated across business lines and legal entities.

Fair Information Practice Principles (FIPPS) are guidelines established by the Federal Privacy Council, originally developed in 1973 to address concerns about data privacy and the use of personal data, particularly in automated systems [18]. FIPPS emphasises accuracy, relevance, timeliness, and completeness, with a strong focus on confidentiality and compliance dimensions related to privacy protection. These principles have influenced numerous privacy regulations and data protection frameworks worldwide.

The European Statistical System Quality Assurance Framework (ESS QAF) provides quality principles and guidelines for European statistical authorities [19]. This framework emphasises accessibility, clarity, relevance, timeliness, accuracy, credibility, efficiency, and consistency, supporting the production of official statistics that meet users' needs while maintaining public trust.

The UK Government Data Quality Framework, shown in Figure 6, is based on the Data Management Body of Knowledge (DAMA DMBOK) and provides guidance for UK public sector organisations [20] [21] [13]. This framework addresses accuracy, completeness, consistency, timeliness, validity, credibility, traceability, and understandability, with particular emphasis on governance and usefulness for government decision-making processes.



Figure 6: Data lifecycle as outlined by the UK Government Data Quality Framework. It describes “the different stages the data will go through from design and collection to dissemination and archival/destruction”.

The International Monetary Fund (IMF) Data Quality Assessment Framework (DQAF) helps countries check and improve the accuracy of their macroeconomic and financial data [22]. The framework includes requirements for quality, methodological soundness, accuracy, and reliability, as well as serviceability and accessibility. It focusses on credibility, consistency, traceability, and understandability in the context of national statistical systems.

The World Health Organization (WHO) Data Quality Assurance (DQA) framework addresses health information systems in global health contexts [23]. This framework emphasises the completeness, timeliness, internal and external consistency, accuracy, and governance of population health data, recognising the unique challenges of health data collection in diverse settings with varying resource constraints.

Framework coverage and selection

The frameworks discussed above vary considerably in their coverage of data quality dimensions, reflecting their different origins, purposes, and domains of application. Figure 7 presents a comparative view of dimension coverage across these frameworks, mapping each framework's dimensions to a standardised vocabulary [2].

This comparison highlights two practical considerations for organisations selecting or adapting a framework. First, core dimensions such as accuracy, completeness,

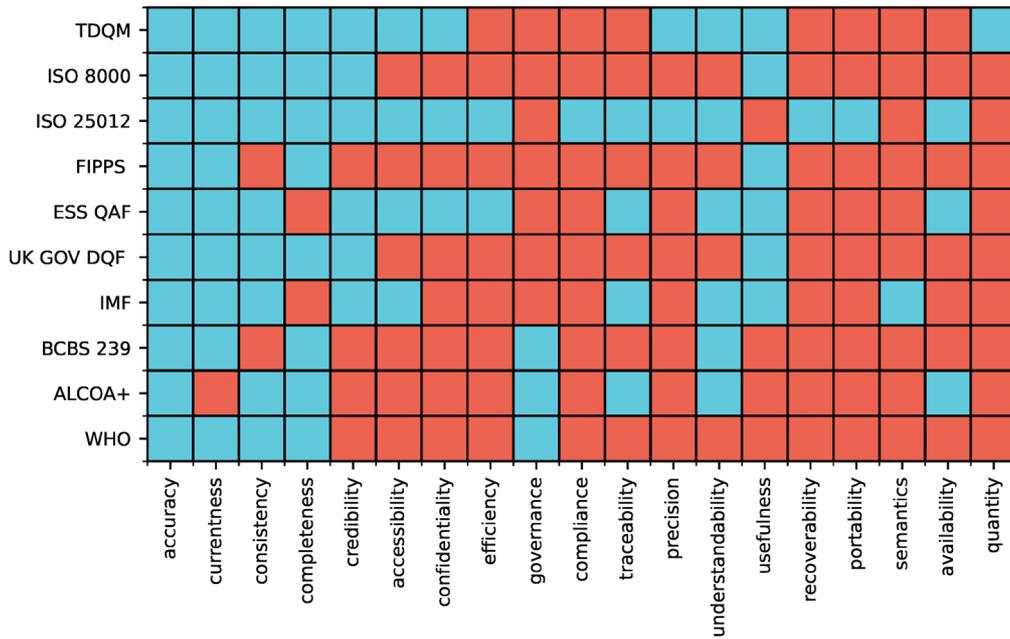


Figure 7: Coverage of data quality dimensions across regulatory-backed frameworks. Blue cells indicate that a dimension is explicitly addressed by the framework; red cells indicate absence. Core dimensions such as accuracy, completeness, and consistency are consistently represented, while dimensions such as semantics, quantity, and portability are addressed by fewer frameworks.

consistency, and currentness are consistently represented across most frameworks, making them a reliable starting point for any data quality programme. Second, organisations should assess whether their specific context requires dimensions that may be absent from their chosen framework. For example, *governance* for regulated environments, or *semantics* for knowledge graph applications.

Generalisable frameworks such as ISO 25012 provide comprehensive coverage suitable for diverse contexts and can serve as foundations for customised approaches. Domain-specific frameworks offer targeted guidance for regulated industries but may require additional dimensions for comprehensive quality assessment. Organisations operating in multiple domains or facing diverse data quality challenges may benefit from hybrid approaches that combine elements from multiple frameworks while maintaining common terminology for consistent communication.

The 19-dimension framework presented in this guide builds upon ISO 25012 by incorporating governance, usefulness, quantity, and semantics as additional data quality

dimensions [1] [2]. This extension provides comprehensive coverage while maintaining the inherent, contextual, and system-dependent classification structure that supports systematic data quality assessment.

Machine-readable data quality

To enable interoperability and machine-readable quality specifications, the data quality dimensions presented thus far have been aligned with the W3C Data Quality Vocabulary (DQV) [9]. This alignment maps dimension definitions presented in this document – aligned with ISO 25012 – as instances within the DQV meta-model, establishing formal relationships between categories, dimensions, metrics, and quality measurements. Figure 8 illustrates this semantic structure, and the full integration mapping is available online².

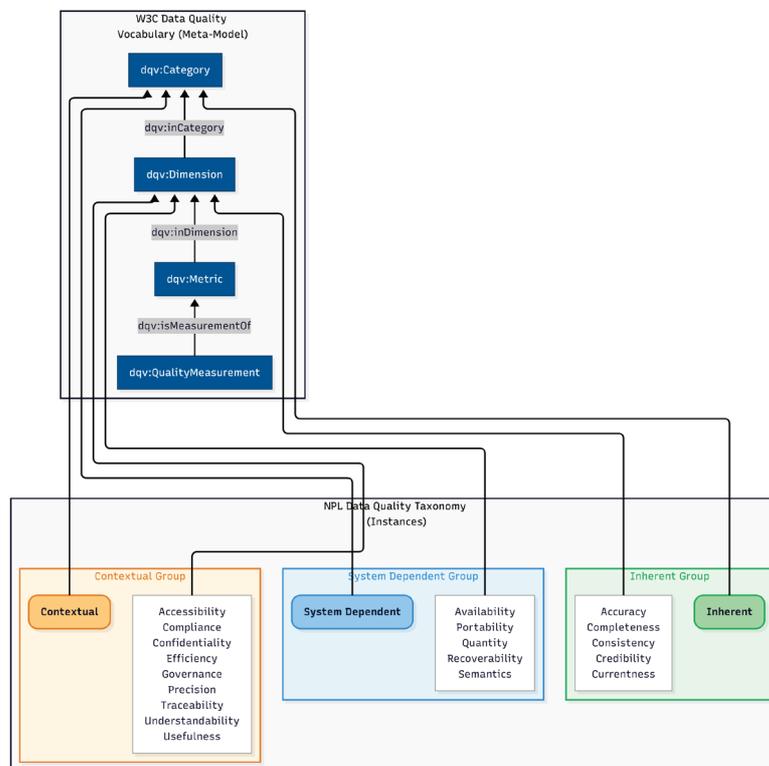


Figure 8: The DQV meta-model (left) defines a hierarchy for describing data quality: Quality Measurements are linked to Metrics, which belong to Dimensions, which are organised into Categories. The NPL Data Quality Taxonomy dimensions (right) are instantiated within this framework, with the three categories mapped as instances of `Category`, and the dimensions mapped as instances of `Dimension`.

² DQV-NPL Integration Mapping: <https://dq-taxonomy-integration-to-dqv-4875bb.gitlab-docs.npl.co.uk/>. This is a live resource under active development; if inaccessible, please contact the responsible author.

This formal structure provides immediate practical benefits. Practitioners can reference standardised dimension definitions when specifying data quality requirements, ensuring consistency across projects and organisations. The alignment with DQV enables interoperability with other quality frameworks and tools that adopt the W3C standard. Additionally, the semantic representation supports automated processing, allowing quality specifications to be validated, compared, and integrated into data management workflows programmatically.

Case Studies

This section highlights how data quality dimensions are interpreted and applied across different contexts. Each case study represents an NPL work that has either been published, been submitted for publication, or is being prepared for submission. They show that the underlying dimensions remain consistent, but their practical interpretation varies depending on the application domain.

Digital verification and validation in manufacturing

Context

Replacing physical testing with digital simulations offers significant cost and time savings in manufacturing, but the credibility of simulation outputs often limits adoption. This case study examined the data requirements for quantifying simulation quality in the remanufacturing of industrial parts, where reconditioned components must demonstrate fitness for redeployment [24].

Approach

As discussed in the Section on “Data Quality Frameworks”, the frameworks can be built upon existing standards and regulations or specifically designed for tailored applications. This case study exemplifies the latter approach: a credibility assessment framework was developed based on NASA's Standard for Models and Simulations (NASA-STD-7009A), adapting its principles to the specific requirements of industrial part remanufacturing [25] [26] [27]. The framework scores simulation quality across eight axes: verification, validation, input pedigree, results uncertainty, results robustness, use history, model and simulation management, and technical review.

A conceptual data model was designed using competency questions to map the data requirements necessary for calculating credibility scores. Three data quality dimensions were of primary relevance. Traceability was interpreted through the data provenance for simulation inputs: whether parameters could be traced back to their original sources and documentation (“Traceability” Subsection). This interpretation aligns with the dimension's emphasis on providing audit trails, applied here to input lineage rather than modification

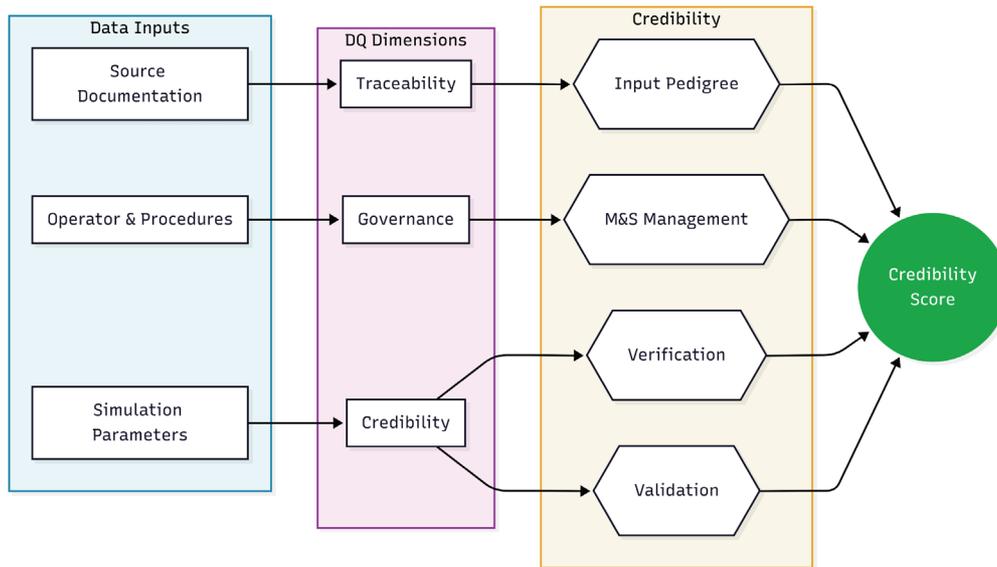


Figure 9: Mapping of data inputs to credibility assessment through data quality dimensions. Source documentation, simulation parameters, and operational procedures are evaluated against relevant data quality dimensions, which in turn inform the credibility axes used to determine whether a reconditioned component is suitable for redeployment.

history. Governance manifested through model and simulation management, which assesses whether simulation procedures follow documented protocols and whether authority and authorisation policies designate who can perform specific data actions (“Governance” Subsection). Credibility was assessed through the trustworthiness of input sources and methodologies, directly reflecting the dimension's definition as the degree to which data are regarded as true and believable by users (“Credibility” Subsection). Figure 9 illustrates how data inputs map to credibility scoring through these dimensions.

Findings

The framework revealed that lower credibility scores were consistently associated with gaps in traceable data. This reinforces the inherent relationship between traceability and credibility discussed in the Traceability subsection. For instance, achieving higher verification scores required not only that simulations produce valid outputs but also that boundary conditions and sensitivity parameters could be traced to documented sources. The data model demonstrated that credibility assessment could be automated when data

were structured to answer specific competency questions about provenance, authority, and documentation.

Implications

This case study illustrates how organisations can adapt generalisable standards to domain-specific needs while maintaining alignment with established data quality dimensions. The interpretation of traceability as provenance documentation, rather than audit trails of data modification, highlights the contextual nature of data quality discussed in the “An Introduction to Data Quality” Section. The interplay between governance and credibility dimensions further highlights that robust data quality practices require attention to multiple dimensions simultaneously, as structured authority and accountability frameworks directly influence the perceived trustworthiness of outputs.

Digital threads in manufacturing

Context

Digital supply chains require the traceable and unaltered flow of data across organisational boundaries, yet incompatibilities between different product lifecycle management systems can compromise data integrity. This case study examined methods for validating data integrity when technical data packs – data packages containing design and manufacturing information – are exchanged between collaborating organisations in a manufacturing supply chain.

Approach

A hashing algorithm tool was developed and validated within the Connected Model-Based Enterprise Environment testbed, which simulated a real-world merge of two digital supply chains [28]. The tool generates unique hash keys from file contents, enabling detection of any modifications, whether accidental or intentional, during transfer. Test cases evaluated the algorithm's ability to identify added, missing, modified, moved, and renamed files across both compressed and uncompressed data formats.

This case study demonstrates how system-dependent data quality considerations (“System-dependent dimensions” Subsection) become critical when data move between

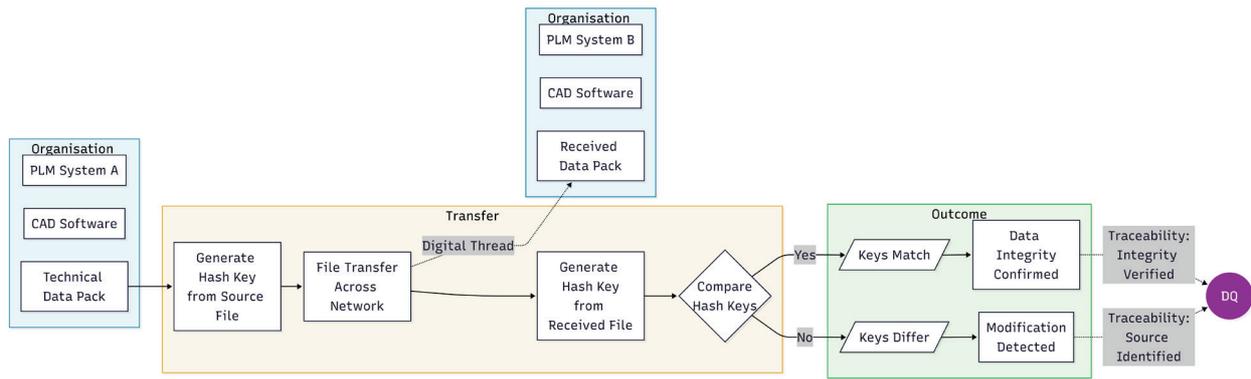


Figure 10: Overview of the hashing-based validation process for digital thread integrity. Technical data packs are transferred between organisations operating different Product Lifecycle Management (PLM) systems. The packs have hash keys generated from file contents at source and destination. Comparison of hash keys confirms data integrity or identifies modifications, supporting the traceability dimension of data quality.

different technological environments. The complex configuration of different technologies in modern data environments requires governance of factors such as portability and recoverability, which are essential not only for reliable analytics but also for operational stability across interconnected systems.

The primary data quality dimension addressed was traceability (“Traceability” Subsection). In this supply chain context, traceability is the ability to verify that data have not been altered during transfer and to identify the source of any changes. This interpretation shifts the focus from documenting who accessed or modified data to confirming that data remain identical to its original state. This subtle but important distinction reflects the operational priorities of supply chain integration. The work highlighted that design intent is primarily linked to file contents rather than metadata, demonstrating how, in this context, accuracy (“Accuracy” Subsection) relates to preserving the original meaning and purpose of data rather than numerical correctness alone. Figure 10 illustrates this validation process.

Findings

The hashing algorithm successfully detected all deliberate modifications across all test cases, regardless of file format or compression status. Implementation in the testbed resulted in a reduction in time resources for validation tasks by automating processes that would otherwise require manual verification. The work confirmed that methods relying

explicitly on file content are more robust than metadata-based validation for maintaining digital thread integrity.

Implications

This case study demonstrates an alternative interpretation of traceability compared to the “Digital verification and validation in manufacturing” case study. Where the digital verification and validation context interpreted traceability as provenance documentation, the digital thread context interprets traceability as integrity verification confirming that data remain unchanged during transfer. Both interpretations are consistent with the definition presented in the “Traceability” Subsection, yet they emphasise different facets of the dimension depending on operational requirements. This highlights the contextual nature of data quality: whether data are considered fit for purpose depends significantly on the application environment and the specific concerns that must be addressed (“An Introduction to Data Quality” Section). Understanding these contextual interpretations enables organisations to implement data quality practices that are aligned with established frameworks and tailored to their specific operational needs.

Electronic health records and clinical trials

Context

Clinical trials depend on the enrolment of cohorts that appropriately reflect the populations who will ultimately receive the intervention. However, traditional approaches for assessing trial representativeness often rely on manually inspecting demographic tables, producing subjective or inconsistent evaluations. Electronic health records (EHRs) are a rich source for generating reference datasets for target populations, but challenges in data fragmentation, incompleteness, and lack of standardised terminology complicate their use. The central question examined in this case study is how the completeness dimension of data quality should be interpreted when EHRs are used to assess clinical trial cohort suitability [29] [30].

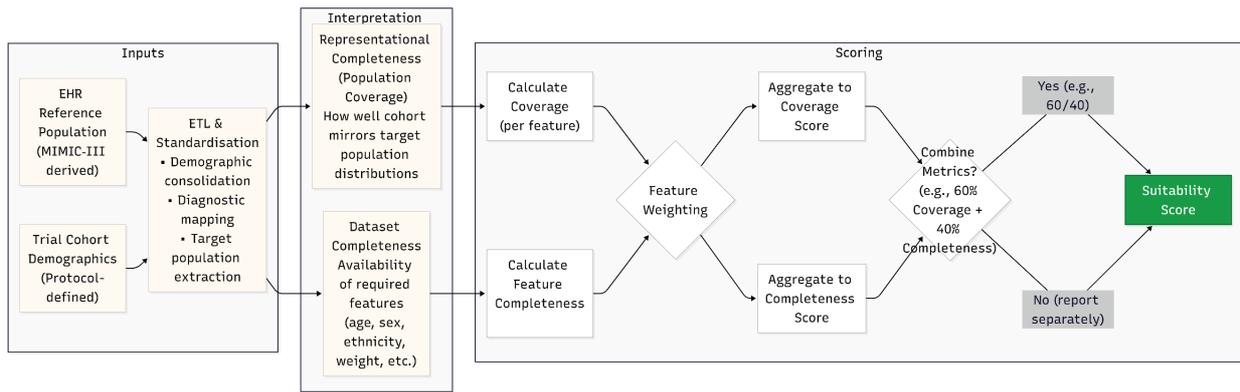


Figure 11: Representativeness assessment workflow showing how dataset completeness and representational completeness are derived from EHR and trial data, weighted, and combined into a suitability score.

Approach

The study developed a quantitative representativeness framework using an EHR-derived population from the MIMIC-III database [31] [32]. A structured ETL pipeline standardised demographic information, mapped diagnostic data to updated terminology, and extracted a disease-specific population to serve as the reference cohort. Representativeness was operationalised through two interpretations of the completeness dimension.

First, coverage completeness quantified how well trial cohorts captured the demographic distribution of the target population across variables such as age, gender, ethnicity, and weight. Second, dataset completeness measured the availability of required demographic attributes for each participant. These metrics were combined into a suitability score.

In this context, completeness extends beyond “presence of data values” (“Generalisable frameworks” Subsection) to include population coverage as an inherent requirement for fitness-for-purpose. This reframing echoes the contextual nature of data quality: dataset incompleteness and population under-representation both indicate incomplete data but differ in origin and impact. Figure 11 illustrates this validation process.

Findings

Demographic alignment was the dominant driver of representativeness. Simulated trial cohorts displayed substantial variation in coverage similarity despite having comparable

completeness levels, confirming the independence of these two completeness components. Datasets with high attribute completeness still received low suitability scores when demographic groups were under-represented. Conversely, even demographically aligned cohorts suffered reductions in suitability when key fields, such as weight or language, showed high dataset incompleteness inherited from the underlying EHR source.

The results also demonstrated that cohort size did not meaningfully influence representativeness, reinforcing that data quality, and not sampling volume, is the primary determinant of suitability. Feature-level analysis showed that high-priority demographic dimensions (age, gender, ethnicity) contributed most to suitability scores, while dataset incompleteness in lower-weight features had limited impact.

Implications

This case study illustrates how traditional data quality dimensions require reinterpretation when applied to certain specific applications. Here, completeness must encompass not only the dataset's internal integrity (missing values) but also its external representativeness (population coverage). This expanded definition aligns with the fitness-for-purpose paradigm: even a fully populated dataset is incomplete if it structurally excludes key demographic subgroups.

This example shows how organisations can integrate metrics based on information theory into clinical trial design workflows, providing reproducible, interpretable, and domain-aligned measures of representativeness. It further exemplifies how the same data quality dimensions can take distinct operational meanings depending on context. In this context, dataset completeness in EHRs versus demographic coverage in trial suitability, while remaining consistent with core definitions within the broader data-quality framework.

Electrocardiogram signal processing and forecasting

Context

ML models are increasingly being considered to support clinical decision, particularly for tasks such as arrhythmia detection from electrocardiogram (ECG) signals. However, ECG data acquisition is highly susceptible to noise, recording artefacts, inconsistent sampling, and variable preprocessing pipelines. These issues threaten the reliability, interpretability,

and safety of downstream ML models. This case study examines how data quality dimensions must be interpreted for timeseries data used in predictive modelling.

Approach

As noted in the “An Introduction to Data Quality” Section, data quality is inherently contextual, meaning that data suitable for one application may be inadequate for another. This case study illustrates how dimensions defined for structured data require reinterpretation when applied to continuous physiological signals.

Using the MIT-BIH Arrhythmia Database on PhysioNet, the study developed a reference-free data quality metric (DQM) constructed from a collection of signal quality indices (SQIs) grounded in ECG physiology [33] [34]. Each SQI captured a specific dimension of data quality. Accuracy ("Accuracy" Subsection) was characterised through measures such as signal-to-noise ratio, sample entropy, and the relative power of the QRS complex – quantifying how faithfully the signal represents true cardiac activity rather than numerical correctness against a reference value. Completeness ("Completeness" Subsection) was assessed by verifying the detectability of the QRS complex, identifying missing values, and evaluating recording continuity, shifting from attribute presence in tabular data to waveform integrity in timeseries signals. Precision ("Precision" Subsection) was interpreted in terms of the consistency and appropriateness of decimal resolution, using indicators such as decimal place variability and the expected resolution derived from ECG device specifications.

Principal Component Analysis (PCA) was used to weight SQIs into a composite metric. Multiple ML models – Convolutional Neural Networks (CNNs), Artificial Neural Networks (ANNs), K-Nearest Neighbours (KNNs), and Random Forests (RFs) – were trained on degraded datasets with systematically reduced accuracy, precision, or completeness. Model performance and calibration were correlated with the DQM to examine how signal quality affects predictive reliability. Figure 12 illustrates this process.

Findings

Data accuracy had the most pronounced influence on model behaviour, reinforcing the inherent nature of this dimension discussed in the "Inherent dimensions" Subsection. As accuracy degraded, models exhibited substantial declines in calibration quality, especially

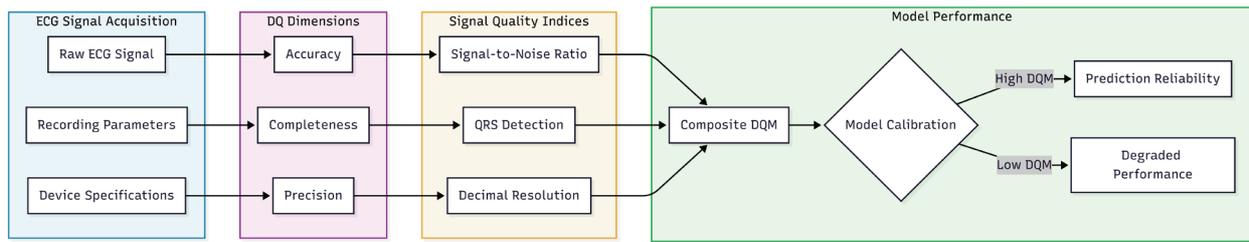


Figure 12: Mapping of ECG signal inputs to model performance through data quality dimensions. Raw signals, recording parameters, and device specifications are evaluated against accuracy, completeness, and precision dimensions via signal quality indices. These combine into a composite data quality metric that informs model calibration and prediction reliability.

in low-confidence prediction ranges. Precision degradations had a lesser effect on DQ, but these effects are consistently negative. Higher DQM scores correlated with higher recall and lower calibration error across models, demonstrating that data quality variance is a meaningful proxy for downstream model robustness.

Model-specific patterns revealed sensitivity differences: CNN and ANN architectures showed strong responsiveness to improved signal quality, while KNN displayed minimal performance variation across DQM levels. Completeness issues, such as missing segments or absent QRS complexes, proved particularly harmful, as they removed physiologically critical information required for correct classification.

Implications

This case study highlighted how data quality dimensions take on domain-specific interpretations when transitioning from table records to physiological signals. Accuracy becomes a measure of physiological fidelity rather than numerical correctness; precision refers to signal-resolution consistency rather than measurement exactness; and completeness is defined by waveform integrity and recording continuity rather than attribute presence. These reinterpretations remain consistent with the definitions presented in the "Data Quality Dimensions" Section, yet they demonstrate the contextual adaptability emphasised throughout this guide.

Comparing this with the manufacturing case studies ("Digital verification and validation in manufacturing" and "Digital threads in manufacturing" Subsections), a common pattern

emerges: the same core dimensions (accuracy, completeness, precision, traceability) manifest differently depending on whether the concern is simulation credibility, supply chain integrity, or clinical signal fidelity. This reinforces the value of a unified framework that provides consistent definitions while accommodating domain-specific interpretation, as discussed in the "Framework coverage and selection" Section.

The findings also demonstrate that a composite data quality metric can act as an effective tool for forecasting model reliability, guiding data collection standards, and informing model-deployment decisions in clinical environments.

AIS data for autonomous vessel operations

Context

Maritime autonomy is advancing rapidly, with Remote Operating Centres (ROCs) managing uncrewed surface vessels and remotely operated vehicles. These centres rely on real time telemetry and sensor data to maintain situational awareness, monitor operations, and intervene when required. However, maritime communications remain challenging due to low bandwidth, high latency, and weather-related signal degradation.

To examine how data quality metrics can quantify these communication issues, we consider a case study using Automatic Identification System (AIS) messages [35]. AIS provides key vessel information including position, Maritime Mobile Service Identity (MMSI), course, and speed, which supports traffic awareness and vessel tracking. This study aimed to assess how reliably such information reaches ROCs under real world connectivity constraints.

Approach

As discussed in the "Data Quality Frameworks" Section, domain-specific frameworks address regulatory and operational requirements within particular sectors. This case study demonstrates how the International Maritime Organization (IMO) Resolution A.1106(29) standard serves as the regulatory basis for assessing AIS transmission quality, highlighting how international standards can underpin data quality assessment in specialised domains [36].

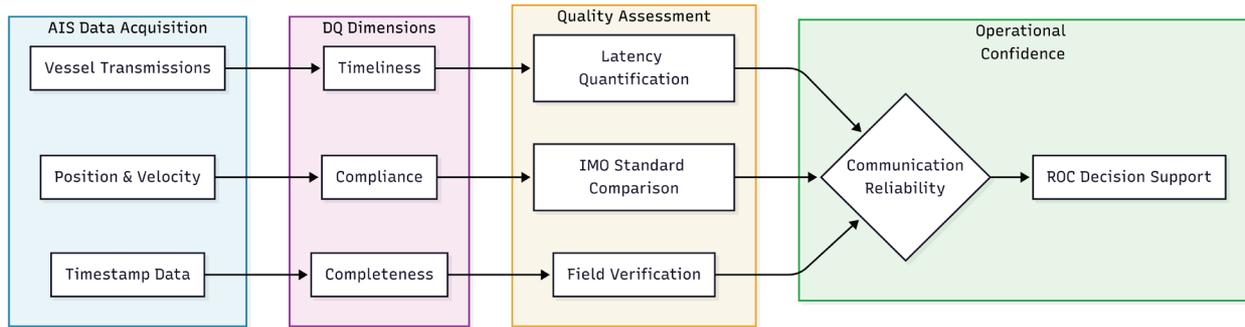


Figure 13: Mapping of AIS data inputs to operational confidence through data quality dimensions. Vessel transmissions, position and velocity data, and timestamps are evaluated against timeliness, compliance, and completeness dimensions, informing communication reliability assessment for Remote Operating Centre decision support.

Two data DQ dimensions were of primary relevance. Timeliness was interpreted as the delay between a vessel's actual state change and when that information reaches the ROC ("Currentness" Subsection). In this context, latency directly impacts the effectiveness of remote monitoring and control. Multiple quantification methods were presented, including quantile-based thresholds, fitting distributions and clustering approaches, to characterise latency patterns across geographic and temporal dimensions.

Compliance was assessed against transmission intervals, which vary according to vessel speed, class, and navigation status ("Compliance" Subsection). This dimension also incorporated completeness, verifying the presence of mandatory fields including vessel identification, velocity, position coordinates, and timestamp in each transmission ("Completeness" Subsection). The process is illustrated in Figure 13.

Findings

Analysis showed that a small percentage of messages were undecodable on average, representing a baseline completeness issue. Geographic analysis showed a positive relationship between distance from the base station and latency, with specific coastal regions showing higher latency rates. Traffic density analysis did not reveal consistent correlations with communication performance, suggesting that geographic factors may be more influential than vessel concentration in this context.

The multiple quantification methods produced consistent spatial patterns despite different threshold calculations, providing confidence that the identified latency trends reflect genuine communication characteristics rather than methodological artefacts.

Implications

This case study demonstrates how data quality dimensions require domain-specific interpretation in maritime communications contexts. Currentness shifts from a general notion of data currency to a precise measure of transmission latency against regulatory intervals. Compliance becomes directly tied to international maritime standards rather than generic data governance requirements. These interpretations remain consistent with the definitions in the "Data Quality Dimensions" Section while addressing the operational priorities of autonomous vessel management.

The work also illustrates the value of applying multiple assessment methods to the same dimension, as discussed in the "Framework coverage and selection" Section, with different quantification approaches providing complementary insights and increased confidence in findings. For ROCs and autonomous vessel operators, systematic data quality assessment provides objective metrics for evaluating communication system performance, supporting both operational safety and regulatory compliance.

Ambient air quality monitoring in metals networks

Context

As discussed in the "Data Quality Frameworks" Section, domain-specific frameworks address regulatory and operational requirements within particular sectors. The UK Metals Network monitors concentrations of selected metals in ambient air at fixed sites distributed across the United Kingdom. Measurement data are uploaded to the UK-AIR platform operated by the Department for Environment, Food & Rural Affairs and are summarised annually by NPL [37] [38]. The regulatory environment governing this work is unique, among the case studies presented thus far, in that data quality expectations are explicitly defined in UK legislation through the Air Quality Standards Regulations, which reference EU Directives 2004/107/EC and 2008/50/EC [39] [40] [41] [42]. These directives prescribe binding data quality objectives for pollutants, including minimum data capture thresholds,

uncertainty limits, and requirements for spatial representativeness based on zone population and pollutant concentration thresholds. This setting therefore provides a clear example of how externally mandated data quality requirements shape the design of measurement systems, influence analytical processes, and determine whether environmental-monitoring data are deemed fit for official reporting.

Approach

Ambient air metals monitoring involves sampling particulate matter (specifically Particulate Matter with diameter of 10 micrometres or less – PM₁₀) using fixed measurement stations that meet siting and device criteria defined in legislation. The network measures four metals required under European directives (nickel, arsenic, cadmium, and lead) and supplements these with additional metals measured by NPL. Each sampling point must generate valid measurements throughout the intended operating period, with downtime accounted for only in cases of scheduled servicing or audits.

The operational workflow integrates several data quality dimensions defined by the legislative framework. Completeness (“Completeness” Subsection) is interpreted as the annual data-capture requirement: a minimum of 90 % valid sampling time per year for each pollutant. Time-related coverage, a sub-dimension of completeness in this context, ensures that data represent the full monitoring period rather than irregular or intermittent sampling. Precision (“Precision” Subsection) is formalised through mandatory uncertainty limits, which set maximum allowable measurement uncertainty (e.g., 40 % for arsenic, cadmium, and nickel; 25% for lead). Spatial completeness is reflected in requirements for the number of sampling points in each zone, determined by population size and whether pollutant concentrations exceed specified assessment thresholds (Upper Assessment Threshold (UAT) and Lower Assessment Threshold (LAT)). This requirement echoes the expanded interpretation of completeness seen in the “Electronic health records and clinical trials” case study, where completeness encompasses representativeness rather than solely data availability.

The network’s quality plan merges these objectives into its operational procedures and informs the software systems responsible for data storage, validation, metadata tracking, and automatic assessment of compliance. This process is shown in Figure 14.

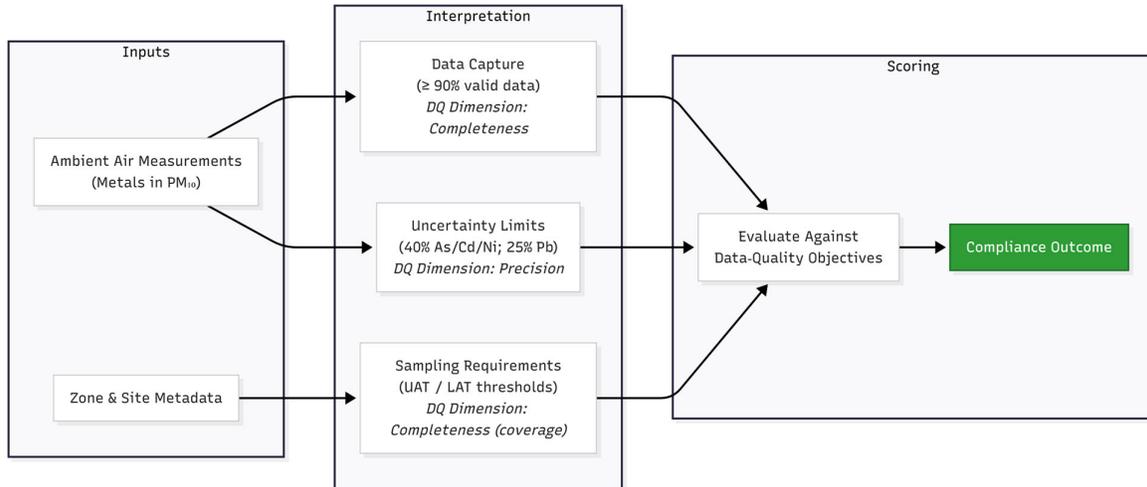


Figure 14: Workflow showing how ambient air metal-concentration measurements, sampling-site information, and operational schedules are interpreted through legally defined data-quality objectives – Including data-capture requirements, uncertainty limits, and spatial coverage thresholds based on UAT and LAT – to determine compliance with regulatory monitoring standards.

Findings

Analysis of data from the UK Metals Network shows that compliance with data quality objectives strongly influences both data availability and interpretability. Sites that fail to meet the 90% data-capture requirement cannot be used for regulatory assessment, regardless of measurement accuracy, reinforcing the inherent relationship between completeness and fitness for purpose discussed in the “Inherent dimensions” Subsection. Similarly, uncertainty limits constrain the choice of instrumentation and calibration procedures. The system's structure ensures that the combination of temporal coverage, spatial distribution, and measurement precision collectively determines whether reported pollutant concentrations are valid for assessing public-health risks.

The inclusion of legally defined thresholds such as UAT and LAT directly affects network design: high-pollution zones require denser measurement coverage, whereas zones below the lower threshold may require fewer sampling points. This creates a data quality framework that is not merely technical but also strategic.

Implications

This case study demonstrates how data-quality dimensions can be codified into legislation, creating objective performance targets that measurement networks must satisfy. The UK Metals Network provides a context where completeness and precision are not abstract concepts, but quantifiable, regulated requirements tied to environmental policy.

The relationship between completeness and fitness-for-purpose is especially clear: even highly precise measurements are disqualified if data capture falls below mandated thresholds, highlighting how incompleteness can undermine regulatory validity. Uncertainty limits the definition of precision as a compliance criterion, ensuring that only measurements falling within acceptable uncertainty bounds are included in official reporting. Spatial coverage requirements illustrate that completeness extends beyond temporal considerations to the geographic dimension, reinforcing the importance of representativeness when monitoring diffuse environmental hazards.

Overall, the UK Metals Network provides an example of how data quality objectives can be operationalised, monitored, and enforced within a national measurement infrastructure, making it a suitable model for illustrating the integration of data quality frameworks in real-world regulatory systems. This creates a parallel to the population coverage interpretation in the clinical trials case study in the “Electronic health records and clinical trials” Subsection.

Comparing this with the other case studies, the UK Metals Network is distinctive in that data quality requirements are externally imposed through legislation rather than developed internally. This demonstrates that the framework presented in this guide (“Framework coverage and selection” Section) can accommodate both organisationally defined and regulatory mandated data quality objectives, providing flexibility for diverse operational contexts.

Conclusions

This guide has presented a framework for understanding and assessing data quality, grounded in established standards and extended to meet the practical needs of data practitioners across diverse domains.

Data quality is inherently contextual. The same dimension may require different interpretation depending on the application environment: simulation credibility assessment interprets traceability as provenance documentation; supply chain integration interprets it as integrity verification; clinical trial design interprets completeness as population coverage rather than attribute presence. Recognising this contextual variation is essential for implementing effective data quality practices.

No single data quality framework suits all contexts. Generalized frameworks such as ISO 25012 provide broad coverage and consistent terminology, making them suitable foundations for most organisations. Domain-specific frameworks such as ALCOA+ or BCBS 239 offer targeted guidance for regulated sectors but may require augmentation when applied elsewhere. The 19-dimension framework presented in this guide aims to balance breadth with practical applicability.

The inherent dimensions (accuracy, completeness, consistency, credibility, and currentness) represent intrinsic data properties and should be supported through robust validation, governance oversight, and thorough documentation. Contextual dimensions (accessibility, compliance, confidentiality, efficiency, governance, precision, traceability, understandability, usefulness) must be evaluated according to user needs and operational constraints. System-dependent dimensions (availability, portability, quantity, recovery, and semantics) require an understanding of the technological and organisational environments in which data operate.

Effective data quality management depends on attention to multiple dimensions simultaneously. The case studies demonstrate that deficiencies in one dimension, such as traceability, can directly affect others, such as credibility. Organisations should prioritise core dimensions (accuracy, completeness, consistency, currentness) as a foundation and then assess which additional dimensions their specific context requires. Regulatory environments may codify certain dimensions into compliance requirements, as illustrated

by the “Ambient air quality monitoring in metals networks” case study, where completeness and precision thresholds are legislatively mandated.

Data quality assessment benefits from complementary methods. Quantitative metrics provide objective measurements where applicable, while qualitative evaluation supports dimensions less suited to numerical scoring. Composite metrics, as shown in the “Electrocardiogram signal processing and forecasting” case study, can act as proxies for downstream system performance. Regardless of the methods used, consistent terminology is essential for comparing assessments and communicating results.

Future Work

The concept of interpretation has been a recurring theme throughout this guide. Each case study has shown that data quality dimensions, while consistently defined, take on different practical meanings depending on the context of application. Traceability in simulation credibility assessment emphasises provenance documentation; in supply chain integration, it emphasises integrity verification. Completeness in clinical trials encompasses population coverage; in environmental monitoring, it encompasses temporal and spatial capture thresholds. This pattern is not incidental; it reflects a fundamental characteristic of data quality assessment that warrants more formal treatment.

The relationship between a dimension and its practical operationalisation can be captured as a structured statement: a data quality dimension is interpreted as a specific measurement or quantification method within a given domain context. This structure, linking a subject to an object through a defined relationship, aligns with the concept of a semantic triple used in knowledge representation and W3C semantic web standards. Figure 15 illustrates this approach. In practice, the same dimension acquires different operational meanings across domains: for instance, traceability may involve hash-based integrity verification in digital supply chains but focus on provenance documentation within simulation credibility assessment. Likewise, completeness may refer to the proportion of valid sampling time in environmental monitoring, while in clinical trial representativeness it relates to the similarity of demographic coverage.

Formalising this interpretive relationship using semantic web standards would offer several benefits. A lightweight ontology pattern could provide a template that organisations extend with their own interpretations, specifying the formal dimension, domain-specific operationalisation, application context, quantification method, and a reference to supporting guidance. Such a framework would support interoperability by enabling organisations operating across multiple domains to navigate between different interpretations of the same dimension. It would facilitate knowledge sharing by creating a repository of validated interpretations that practitioners can consult when designing assessments for new applications. It would enhance regulatory alignment by providing a structured way to document how legislatively mandated requirements map to established

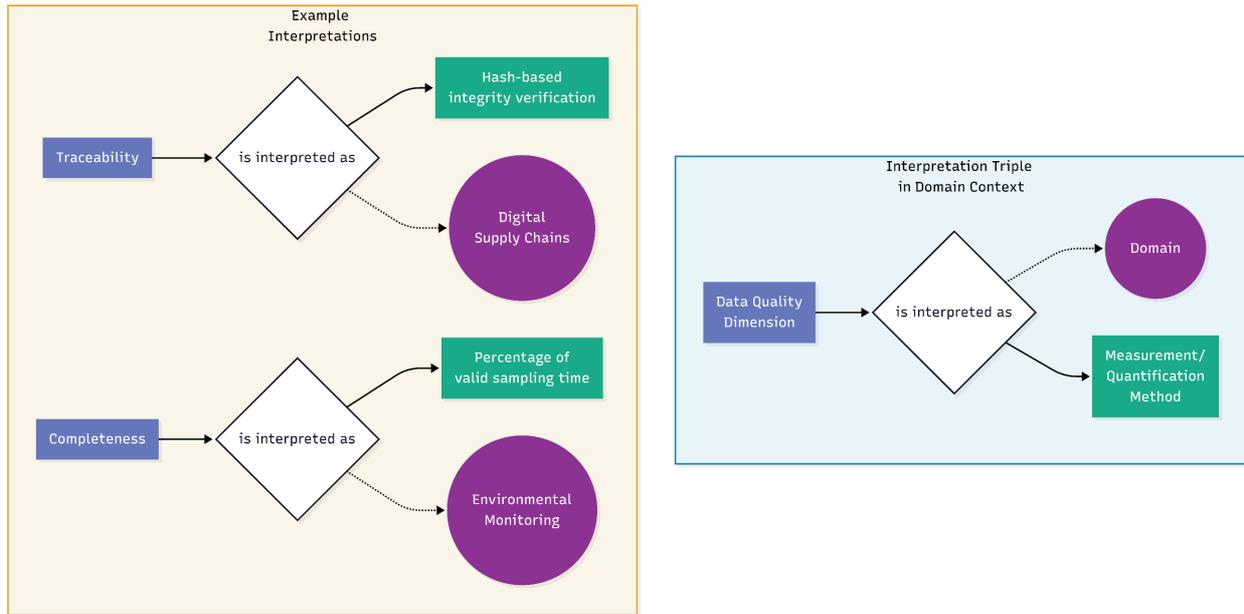


Figure 15: Example of semantic triple structure for data quality dimension interpretation. Domain-specific examples (left) and “blank” template example (right). A data quality dimension (purple rectangles) is linked to its measurement or quantification method (green rectangles) through an "is interpreted as" relationship (white diamond shapes), with the domain (purple circles) providing the contextual frame within which the interpretation applies.

dimensions. It would also offer a machine-readable foundation for assessment tools that adapt their metrics to different domain contexts.

The vision is a living framework that grows as practitioners contribute interpretations from their respective domains. Rather than prescribing how dimensions should be measured, such a resource would capture how they are measured across the data quality community, creating a collective vocabulary that bridges the gap between abstract definitions and concrete assessment practices. This guide, with its emphasis on contextual interpretation, represents a meaningful step towards that goal.

Acknowledgements

This work was funded by the UK Government Department for Science, Innovation and Technology through the UK's National Measurement System. Thanks also to Hannah Strassburg and Chris Barnes for providing feedback on the original manuscript.

References

- [1] R. Miller, H. Whelan, M. Chrubasik, D. Whittaker, P. Duncan and J. Gregório, "A Framework for Current and New Data Quality Dimensions: An Overview," *Data*, vol. 9, no. 12, p. 151, 2024.
- [2] R. Miller, S. H. M. Chan, H. Whelan and J. Gregório, "A Comparison of Data Quality Frameworks: A Review," *Big Data and Cognitive Computing*, vol. 9, no. 4, p. 93, 2025.
- [3] S. Zahidi, "The Future of Jobs Report 2025," World Economic Forum, Cologny/Geneva, 2025.
- [4] L. Wright, K. Khatry, M. Bevilacqua, L. Wright, M. Chrubasik and J. Gregorio, "Twins, pyramids and environments: unifying approaches to virtual testing," *Engineering Modelling, Analysis and Simulation*, vol. 3, no. 1, 2026.
- [5] L. Wright and S. Davidson, "How to tell the difference between a model and a digital twin," *Advanced Modeling and Simulation in Engineering Sciences*, vol. 7, no. 1, p. 13, 202.
- [6] L. Budach, M. Feuerpfeil, N. Ihde, A. Nathansen, N. Noack, H. Patzlaff, F. Naumann and H. Harmouch, "The effects of data quality on machine learning performance on tabular data," *Information Systems*, vol. 132, p. 102549, 2025.
- [7] International Organization for Standardization, "ISO 8000-1:2022 Data Quality—Part 1: Overview," 2022. [Online].
- [8] International Organization for Standardization, "ISO/IEC 25012:2008 Software Engineering—Software Product Quality Requirements and Evaluation (SQuaRE)—Data Quality Model," 2008. [Online].

- [9] W3C Working Group, "Data on the Web Best Practices: Data Quality Vocabulary," W3C, 2016. [Online]. Available: <https://www.w3.org/TR/vocab-dqv/>. [Accessed 9 February 2026].
- [10] Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1), "Guide to the expression of uncertainty in measurement," Bureau International des Poids et Mesures (BIPM), 2023.
- [11] Working Group 2 of the Joint Committee for Guides in Metrology (JCGM/WG 2), "International Vocabulary of Metrology," Bureau International des Poids et Mesures (BIPM), 2023.
- [12] L. Bertossi, F. Rizzolo and L. Jiang, "Data quality is context dependent," in *International Workshop on Business Intelligence for the Real-Time Enterprise*, Berlin, Heidelberg, 2010.
- [13] DAMA International, DAMA-DMBOK Data Management Body of Knowledge, 2nd ed., Sedona, AZ, USA: Technics Publications, 2017.
- [14] MIT Information Quality Program, "Total Data Quality Management (TDQM) Program," 2024. [Online]. Available: <http://mitiq.mit.edu/>. [Accessed 14 August 2024].
- [15] International Organization for Standardization, "ISO 9001:2015 Quality Management Systems—Requirements," 2015. [Online].
- [16] M. Durá, A. Sánchez-García, C. Sáez, F. Leal, A. E. Chis, H. González-Vélez and J. M. García-Gómez, "Towards a computational approach for the assessment of compliance of ALCOA+ Principles in pharma industry," in *Challenges of Trustable AI and Added-Value on Health*, Amsterdam, The Netherlands, IOS Press, 2022, pp. 755-759.
- [17] Basel Committee on Banking Supervision, "Principles for Effective Risk Data Aggregation and Risk Reporting," Bank for International Settlements, Basel, Switzerland, 2013.

- [18] Federal Privacy Council, "Fair Information Practice Principles (FIPPS)," 2024. [Online]. Available: <https://www.fpc.gov/>. [Accessed 14 August 2024].
- [19] Eurostat, "European Statistics Code of Practice—Revised Edition 2017," Publications Office of the European Union, Luxembourg, 2018.
- [20] Government Data Quality Hub, "The Government Data Quality Framework," 2020. [Online]. Available: <https://www.gov.uk/government/organisations/government-data-quality-hub>. [Accessed 1 October 2024].
- [21] Government Data Quality Hub, "The Government Data Quality Framework: Guidance," 2020. [Online]. Available: <https://www.gov.uk/government/publications/the-government-data-quality-framework/the-government-data-quality-framework-guidance>. [Accessed 1 October 2024].
- [22] International Monetary Fund, "Data Quality Assessment Framework (DQAF)," 2003. [Online]. Available: <https://www.imf.org/external/np/sta/dsbb/2003/eng/dqaf.htm>. [Accessed 1 October 2024].
- [23] World Health Organization, "Data Quality Assurance: Module 1: Framework and Metrics," World Health Organization, Geneva, Switzerland, 2022.
- [24] J. Gregório, M. Alsuleman, M. Chrubasik, P. Duncan and G. Bisland, "A competency question driven approach to conceptual data model design for digital verification and validation," in *Advanced Mathematical and Computational Tools in Metrology and Testing XIII*, Sarajevo, 2025.
- [25] M. J. Steele, "Standard for models and simulations," NASA, 2016.
- [26] M. J. Steele, "The NASA standard for models and simulations," in *Proceedings of the 2007 Summer Computer Simulation Conference*, 2007.

- [27] M. Alsuleman, A. Almudahka, R. Miller, M. Chrubasik, P. Duncan and J. Gregório, "Enhancing simulation credibility through data model development: A case study using NASA-STD-7009," *Measurement: Sensors*, p. 101461, 2024.
- [28] J. Gregório, R. Miller and P. Duncan, "Hashing it out: Unravelling the threads of data in digital manufacturing," *Measurement: Sensors*, p. 101802, 2024.
- [29] J. Gregório, A. Lemanska, B. Cieszynski, N. Perić, M. Alsuleman, M. Chrubasik and P. Duncan, "An EHR Data Standardisation Pipeline Using the MIMIC-III Dataset: Foundation for a Clinical Trial Data Quality Assessment," in *Proceedings of the 19th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 3: HEALTHINF*, 2026.
- [30] J. Gregório, A. Lemanska, B. Cieszynski, N. Perić, M. Alsuleman, M. Chrubasik and P. Duncan, "Beyond Demographic Tables: Integrating Data Quality in Clinical Trial Representativeness," *[Submitted for Publication]*, 2025.
- [31] A. E. W. Johnson, T. J. Pollard, L. Shen, L.-w. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi and R. G. Mark, "MIMIC-III, a freely accessible critical care database," *Scientific Data*, vol. 3, no. 1, p. 160035, 2016.
- [32] A. Johnson, T. Pollard and R. Mark, "MIMIC-III Clinical Database (version 1.4)," 2016. [Online]. Available: <https://doi.org/10.13026/C2XW26>.
- [33] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215-e220, 2000.
- [34] B. Cieszynski, P. Krishnadas, V. Desai, N. Perić, M. Alsuleman, M. Chrubasik, P. Duncan and J. Gregório, "Multi-Dimensional Data Quality Metrics for Machine Learning in Arrhythmia Classification," *[Submitted for Publication]*, 2025.

- [35] S. H. M. Chan, J. Gregório, M. Bevilacqua, J. MacDonald, J. Pewter, J. Fishwick and R. Miller, "Assessing AIS Data Timeliness: A Proof of Concept for Autonomous Vessel Operations," *[in Preparation]*, 2025.
- [36] J.-U. Schröder-Hinrichs, E. Hollnagel, M. Baldauf, S. Hofmann and A. Kataria, "Maritime human factors and IMO policy," *Maritime Policy & Management*, vol. 40, no. 3, pp. 243-260, 2013.
- [37] Department for Environment, Food & Rural Affairs, "Heavy Metals Network," 2024. [Online]. Available: <https://uk-air.defra.gov.uk/networks/network-info?view=metals>. [Accessed 9 February 2026].
- [38] K. R. Williams, E. C. Braysher, A. S. Brown, R. J. C. Brown, D. M. Butterfield, J. H. L. Cheong, C. C. Robins and Y. A. Showman, "The UK Heavy Metals Monitoring Network: Annual Report 2024," National Physical Laboratory, 2025.
- [39] European Parliament and Council of the European Union, "Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European Statistics," European Union, Brussels, Belgium, 2009.
- [40] European Parliament and Council of the European Union, "Directive 2004/107/EC of the European Parliament and of the Council of 15 December 2004 relating to arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air," 2004. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32004L0107>. [Accessed 9 February 2026].
- [41] European Parliament and Council of the European Union, "Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe," 2008. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32008L0050>. [Accessed 9 February 2026].
- [42] UK Government, "The Air Quality Standards Regulations 2010," 2010. [Online]. Available: <https://www.legislation.gov.uk/ukxi/2010/1001/contents>. [Accessed 9 February 2026].

Data Quality Good Practice Guide

Providing the measurement capability
that underpins the UK's prosperity
and quality of life

To find out more about NPL:

npl.co.uk

To get in touch:

npl.co.uk/contact

020 8977 3222

Hampton Road

Teddington

Middlesex

TW11 0LW

**Please consider the environment
before printing this document**